THE GOVERNANCE OF

INFORMATION TECHNOLOGY SERVICE PROVISION

By

Nikolaos A. Mylonopoulos

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Warwick Business School

University of Warwick

TABLE OF CONTENTS

ABSTRACT.....vii

DECLARATION.....viii

ACKNOWLEDGEMENTS.....ix

CHAPTER ONE

FROM '*OUTSOURCING VERSUS INSOURCING*' TO THE GOVERNANCE OF INFORMATION TECHNOLOGY SERVICES.....1

1. INTRODUCTION.....1

2. TOWARDS A DEFINITION OF OUTSOURCING.....7

3. DRIVERS, EXPECTATIONS BENEFITS AND RISKS.....10

3.1. Financial and economic expectations.....12

3.2. Strategic and organisational benefits and risks.....15

3.3. Technical and technological reasons for outsourcing.....17

3.4.Political antecedents to outsourcing.....18

4. THE EVOLVING CONCEPTUALISATION OF IT OUTSOURCING.....19

5. ECONOMICS IN IT OUTSOURCING RESEARCH.....26

6. THE GOVERNANCE OF THE IT FUNCTION AND OUTSOURCING.....33

7. CONCLUSIONS.....39

CHAPTER TWO

ON THE ECONOMICS OF THE ORGANISATION OF PRODUCTION AND EXCHANGE.....41

1.INTRODUCTION.....41

2. INSTITUTIONAL ENVIRONMENT: THE TRANSACTION AND THE FIRM.....48

2.1. Relational exchange.....49

2.2. The firm as a discrete alternative to the market.....51

3. TRANSACTION COST ECONOMICS.....54

3.1. Behavioural assumptions.....55

3.2. Dimensionalising transactions.....57

3.3. Dimensionalising governance structures 59

3.4. Aligning transactions with governance structures 62

4. AGENCY COSTS 67

4.1. Mathematical principal-agent theory 69

4.2. Positive agency theory 73

5. INCOMPLETE CONTRACTS AND PROPERTY RIGHTS 77

6. MEASUREMENT COSTS AND THE FIRM AS KNOWLEDGE REPOSITORY 81

7. CULTURE, REPUTATION AND TRUST 84

8. CONCLUSIONS 88

8.1. The theoretical basis of this thesis 91

CHAPTER THREE

AN EPISTEMOLOGICAL BASIS FOR RESEARCH ON THE GOVERNANCE OF INFORMATION TECHNOLOGY PROVISION 101

1. INTRODUCTION 101

2. EPISTEMOLOGICAL AND METHODOLOGICAL CONCERNS IN IS RESEARCH 103

3. THE MAIN TENETS AND IMPLICATIONS OF TRANSCENDENTAL REALISM 110

4. THE PROJECT UNDERTAKEN IN THESIS 119

5. INCORPORATING ECONOMICS INTO A REALIST THEORY 121

6. EMPIRICAL METHODS 126

6.1. Realism 127

6.2. Action 131

7. CONCLUSIONS 135

CHAPTER FOUR

THE PROBLEM FACING PRACTITIONERS: A CASE STUDY IN BRITISH PETROLEUM PLC 137

1. INTRODUCTION 137

2. CARRYING OUT THE CASE STUDY 138

2.1. Strengths and limitations of the approach 147

3. ORGANISATIONAL BACKGROUND 149

4. IT PROCUREMENT PHILOSOPHIES IN THE BP GROUP 152

5. BPX SUPPLIER SELECTION 157

6. THE TRANSITION TO OUTSOURCING 159

7. THE STRUCTURE OF CONTRACTS 160

7.1. Pricing and incentives *165*

7.2. Site organisation *173*

8. FURTHER ISSUES OF CONCERN 176

9. CONCLUSIONS 182

CHAPTER FIVE

SERVICE CLUSTERING AND GOVERNANCE ESTABLISHMENT (S-CAGE): BRIDGING THEORY AND PRACTICE 185

1. INTRODUCTION 185

2. OVERVIEW OF S-CAGE 188

3. SERVICE CLUSTERING 194

4. THE GOVERNANCE OF SERVICE CLUSTERS 197

4.1. Assurance of service delivery *203*

4.1.1. Service specification 203

4.1.2. Service level requirements specification 211

4.1.2.1. Lifecycle of service level requirements 214

4.1.3. Service level measurement and verification 220

4.2. Incentives for service level attainment *225*

4.2.1. Implicit incentive patterns 228

4.2.2. Contractual reward/penalty devices 233

4.2.3. Other contractual and organisational safeguards 238

4.3. Organisation and management of service provision *241*

4.3.1. Roles and responsibilities 242

4.3.2. Co-ordination and communication mechanisms 244

4.3.3. Locus and scope of decision making 246

5. CONCLUSIONS 249

CHAPTER SIX

USING S-CAGE TO UNDERSTAND IT OUTSOURCING PRACTICES 252

1. INTRODUCTION 252

2. CORPORATE IT OUTSOURCING PRACTICES AT ICI 254

2.1. Strategic direction 255

2.2. Governance portfolio 257

2.3. Origin contract: Governance structure 259

2.3.1. Assurance of service delivery 259

2.3.1.1. Service specification 259

2.3.1.2. Service level requirements specification 265

2.3.1.3. Service level measurement and reporting 268

2.3.2. Incentives for service level attainment 269

2.3.2.1. Implicit incentive patterns 269

2.3.2.2. Contractual reward/penalty devices 271

2.3.2.3. Other contractual and organisational safeguards 274

2.3.3. Organisation and management of service provision 274

2.3.3.1. Roles and responsibilities 274

2.3.3.2. Coordination and communication mechanisms 275

2.3.3.3. Locus and scope of decision making 276

2.4. Comments 277

3. TOTAL IT OUTSOURCING AT ANGLIAN WATER 280

3.1. Historical background 280

3.2. Governance portfolio 281

3.3. CSC contract: Governance structure 282

3.3.1. Assurance of service delivery 282

3.3.2. Incentives for service level attainment 283

3.3.3. Organisation and management of service provision 285

3.4. Comments 287

4. CONCLUSIONS 290

CHAPTER SEVEN

CONCLUDING REMARKS AND THE WAY AHEAD.....296

1. INTRODUCTION.....296

2. CONTRIBUTION AND LIMITATIONS.....300

2.1. A note on validity.....303

3. FURTHER RESEARCH DIRECTIONS.....305

3.1. Governance feasibility frontiers.....308

3.1.1. Research designs for governance feasibility.....316

4. INFORMATION TECHNOLOGY AND OTHER BUSINESS ACTIVITIES.....318

5. A FINAL ILLUSTRATION.....320

REFERENCES.....324

APPENDIX 1: INTERVIEW GUIDE.....351

APPENDIX 2: RESEARCH PARTICIPANTS.....357

LIST OF FIGURES

CHAPTER TWO

FIGURE 2.1: A classification of the literature on economic organisation 48
FIGURE 2.2: The governance continuum with indicative categories of governance 98
FIGURE 2.3: The governance continuum with indicative IT contracts 99

CHAPTER THREE

FIGURE 3.1: Types of research from a realist philosophical perspective 127

CHAPTER FOUR

FIGURE 4.1: BP Exploration's organisational transition to outsourcing 155
FIGURE 4.2: The grief point system of the BPO/SEMA call off contract 169

CHAPTER FIVE

FIGURE 5.1: The governance continuum with indicative IT contracts 189
FIGURE 5.2: The clustering continuum and the main benefits of each extreme 192
FIGURE 5.3: IT service clustering and governance establishment (S-CAGE) 193
FIGURE 5.4: The governance of IT service clusters as a layered system 200
FIGURE 5.5: Lifecycle of service level requirements 215

CHAPTER SIX

FIGURE 6.1: The organisation of Information Services at Anglian Water 285

CHAPTER SEVEN

FIGURE 7.1: Frequency distributions of governance structures for individual services 310
FIGURE 7.2: Examples of service clusters 311
FIGURE 7.3: Governance portfolio allocation 312
FIGURE 7.4: Governance feasibility frontier 313
FIGURE 7.5: Demand and supply side governance feasibility frontiers 315
FIGURE 7.6: Number of articles on 'information technology outsourcing' in the ABI/INFORM bibliographic database, over the period 1988-1996 321

LIST OF TABLES

CHAPTER THREE

TABLE 3.1: The three domains of the world according to transcendental realism 111

CHAPTER FOUR

TABLE 4.1: Outsourcing criteria, reasons and expectations for BPO and BPX 157

CHAPTER FIVE

TABLE 5.2: The three layers of governance and the corresponding insights from the theories reviewed in chapter 2 202

CHAPTER SEVEN

TABLE 7.1: Governance portfolio allocation (hypothetical example) 312

ABSTRACT

The governance of information technology (IT) service provision entails all those elements of the structure and process of contracting and organisation that are necessary in order to deliver effective and efficient IT services within today's organisations. This thesis develops a framework integrating those elements of governance with a view to explaining actual governance structures and to guiding relevant decisions in practice.

Until the late 1980s, the principal concerns with respect to the organisation of IT activities revolved around the issues of centralisation versus decentralisation and end-user computing versus specialist control. Whilst contracts for various systems and services have always been a significant part of the IT department's activity, the issue of structuring such contracts took a prominent position on the management agenda in the beginning of 1990s with the proliferation of large contracts of the 'total outsourcing' kind. IT outsourcing rekindled interest in and reshaped the agenda of the organisation of IT provision. The debate on IT outsourcing provided the motivation and the starting point of this thesis.

Thus, chapter 1 reviews this debate and, following on the opinion of others, recasts the question of 'whether to outsource or not' as 'how to set up the governance (i.e. the structure and process of contracting and organisation) of IT service provision'.

In order to shed some light on the elements of governance, chapter 2 turns on a range of economic theories of the firm. These theories are critically reviewed and some preliminary suggestions as to how they might inform the governance of IT are put forth. The chapter concludes by setting the theoretical foundations for the rest of the thesis. The notion of a governance continuum between the ideal pure market and the ideal pure hierarchy is introduced. Actual governance structures, it is argued, can be placed on this continuum as individual instances.

Before advancing onto more substantive work, chapter 3 pauses to reflect on the epistemological basis of this research project. The main principles of the epistemological position adopted here are taken from the philosophical arguments

of transcendental realism. The implications of this epistemological position for the empirical methods and the theoretical claims made in this thesis are also examined.

Chapter 4 documents an intensive case study at British Petroleum Plc. The purpose of this case study was to draw lessons from practice and to assess the relevance of the theories of chapter 2. This case study was an interactive learning process through which the researcher sought access to management practice in order to assess economic theories, while the managers at BP sought a broader understanding of IT outsourcing. The outcome of this interaction was the S-CAGE framework which coupled theoretical insights with practical relevance.

The S-CAGE (Service Clustering And Governance Establishment) framework is described in detail in chapter 5. It is put forth as both an explanatory and a normative account of the governance of IT service provision. It is based on the notion of governance continuum, it introduces the idea of grouping services into clusters and it provides a classification of the elements of governance that should be customised to the characteristics of each cluster of services.

Chapter 6 presents two further case studies aimed at evaluating the usefulness of S-CAGE in understanding and explaining alternative outsourcing practices. ICI and Anglian Water have been visited for this purpose. On the basis of these cases, the concluding chapter summarises the strengths and limitations of the proposed framework. An attempt is also made to set forth some preliminary theoretical ideas extending the notions of clustering and governance continuum. The thesis concludes with a final short illustration of the use of the S-CAGE framework.

DECLARATION

The following article is solely based on work that I carried out during the early stages of developing the S-CAGE framework.

Mylonopoulos N.A. and Ormerod R.J., "*A microanalytic approach to the efficient governance of Information Technology service provision: The case of outsourcing*", Warwick Business School Research Paper No. 142, December 1994. Also in the Proceedings of the Third European Conference on Information Systems, Athens, Greece, June 1-3 1995, pp. 749-765.

ACKNOWLEDGEMENTS

The Warwick Business School (WBS) PhD has been a tremendous learning experience for me. The principal figure in this process has been my supervisor, Professor Richard Ormerod. Richard encouraged me from the beginning to embark on an intellectual exploration of various theoretical ideas and philosophical arguments. However, he has a sharp pragmatist eye for identifying the relevant points among a complex set of elaborate theorising. I am wholeheartedly thankful to Richard for his friendly and really useful guidance and continuing support, both on the development of the thesis as well as on many practical matters of the PhD process, including the contacts with the three companies for my case study work.

This thesis relies a lot on economic theories. Professor Mike Waterson, from the Economics department at Warwick, has been a reliable source of constructive criticism and advice throughout these years. Although the methods and style of my work are foreign to the tradition in Economics, he willingly read the majority of my drafts and made many insightful comments and recommendations. Apart from safeguarding the theoretical background of my research, Mike's assistance also helped strengthen my own confidence in handling these theories. I am grateful to him.

The learning experience of the WBS PhD relies on the many leading academics in the School who are always willing to offer advice. Moreover, the vibrant community of fellow PhD students provides a strong sense of belonging and a constant source of friendship and support. Over the years, the many friends I had and the many new friends I made played their own part in helping me survive the PhD.

I owe a lot to Antreas Athanassopoulos. His tough-minded attitude has been an example of commitment to hard work and achievement in all aspects of life. His genuine friendship made hectic schedules enjoyable. Antreas read some of my drafts and generously offered his advice and the hard-earned lessons he learned from his own experience.

Professor George Doukidis has always been encouraging and supporting me, since I was an undergraduate student at the Athens University of Economics and Business. He did not have a direct involvement in my PhD but had a great influence on my academic development. Rachel Allen, Georgia Siora and Ifigenia Pottaki made comments on various drafts by exhaustively scrutinising every single word that I had written. This is the kind of review and feedback every PhD student dreams of but rarely finds.

Dimitris, Sophia, Christina and Dimitra, my cousins, have been a constant source of courage and motivation even if they have been physically located in distant places all over the world. Last, but not least, Alexandros and Eleni, my parents, always stood by me, quietly watched my steps, joined in my joys and sorrows and did everything to support me. There really are no words to express my feelings. I dedicate this thesis to them.

From '*Outsourcing Versus Insourcing*' To The Governance Of Information Technology Services

CHAPTER ONE

1. Introduction

Information technology outsourcing is a contentious issue in both the academic and practitioner communities. It appears to have strong followers as well as fierce opponents as it has demonstrated both successes and failures. This thesis aspires to delve deeper into the management phenomenon and to observe the trends without being carried away by the outsourcing bandwagon. The aim is to make a contribution both to our understanding as well as to the improvement of the management practice of outsourcing

On 28 November 1996 *Computer Weekly* presented news of two completely opposite approaches to IS organisation. David Jones, then managing director of the information systems division of Scottish Power,is quoted saying that

> "*you can make tactical use of outsourcing, say for sunset systems, but the IT strategy must remain internal, as should systems that are competitive ... I can offer* [our internal customers] *BMW services at a Ford Escort price.*"(Vowler 1996)

The second news article is about Ford's £27 million contract with software house Logica. Ford UK, having outsourced various parts oftheir IT activity, assigned the management of their entire IT supply chain toLogica. Moreover, Ford hopes in the long run to appoint a single company to manage its IT supplyworld-wide (Black 1996). It seems that it is not only the technical capability that Ford does not want to maintain in-house but the management and coordination of suppliers as well.

Both companies are large and highly successful in their respective markets. Yet each follows a completely different approach with respect to the use of external vendors for IT services. Moreover, as Valor et al (1993) found, companies outsourcing their IT services display greater diversity of opinion among senior managers with regards to outsourcing, as compared to companies running traditional IT departments. How can one make sense of such contradictory opinions and practices in a dynamic technological and market environment?

Examples of diverse approaches and practices when companies decide to contract out some parts of their IT abound (e.g. Griese 1993, Altinkemer et al 1994, Pinnington and Woolcock 1994, Sobol and Apte 1995). In 1996 ICI announced[1] a five year contract with Origin, worth £75 million. It was believed to be the largest such contract in the chemicals industry up to then. The deal was considered beneficial for both parties, as ICI managed to downsize without resorting to redundancies and Origin used this contract and the skills acquired as a springboard to expand its business in the UK.

When Gloucester County Council in England decided to switch from one facilities management supplier to another, it faced strong resistance from the incumbent supplier's employees (former council staff) whose jobs were threatened. Both the Council and the new supplier hoped that staff would also transfer to the new supplier. They, however, decided to remain with the same employer who, in turn, tried to find jobs for them within the group and denied access to staff by the new supplier who wanted to make job offers to them. The existing staff has company-specific knowledge of the systems and operations and is invaluable both to the

---

[1] "Staff are winners as ICI goes Dutch for FM deal", *Computer Weekly*, February 8, 1996.

council and to the new supplier. This is an interesting case of the costs of switching IT service provider[2].

Rolls-Royce entered into a £300m IT outsourcing contract with EDS. According to the contract, EDS and its consultancy division, AT Kearney, make their own investment in new IT for Rolls-Royce in the expectation that it will lead to efficiency savings at the customer site, resulting in job cuts. If it succeeds, EDS will receive a share of the savings. If it fails, it will have to write off its investment costs[3] (Ulset 1996).

Other examples include RAC which separated mainframe operations and support from software development for the mainframe and assigned each service to a different supplier[4]. The insurance company Sun Alliance transferred 140 staff to IBM as part of another outsourcing deal[5]. The contract of DuPont with CSC and Andersen consulting is considered the largest in the world, worth £2.6 billion and involving more than 4000 IT staff[6].

Meanwhile, it seems that some companies make a call for tenders for the sole purpose of benchmarking their internal operations and getting ideas for improving them, without any serious intention of hiring a vendor. As a consequence, IT suppliers have become wary of such attempts because of the time and effort required to write proposals.

---

[2] Collins T (1996), 'Staff rebel against FM supplier change', *Computer Weekly*, Thursday 5 December, p. 1.

[3] Collins T. (1997), "EDS secures £300m Rolls-Royce power group outsourcing deal", *Computer Weekly*, Thursday 9 January, p. 2.

[4] Collins T (1997), 'Two FM suppliers to get share of RAC business', *Computer Weekly*, Thursday 16 January, p. 6.

[5] "IBM clinches Sun Alliance outsourcing deal", *Computer Weekly*, February 29 1996.

[6] Collins T (1997), 'DuPont gives FM suppliers equal share of $4bn deal', *Computer Weekly*, Thursday 2 January, p. 2.

The public sector is also a very large user of IT outsourcing contracts, at least in the UK (Srowger, 1993; Currie, 1996). As part of the Competing For Quality (CFQ) initiative the UK public sector has outsourced a large part of its IT services, 10% of which is said to have been awarded to EDS[7]. Major government departments such as the Inland Revenue and the Department of Social Security have outsourced the whole of their IT operations to EDS. In September 1996 the press published information about the UK government's plans to outsource some or all IT services of the Ministry of Defence intelligence services, despite the sensitivity and confidentiality of the systems and data[8].

Outsourcing is not new nor is it limited to IT. It is being widely adopted in various business domains such as project management, accounting, manufacturing, logistics and distribution (Szymankiewicz 1993, Ward and Chapman 1994, Rothery and Robertson 1995). In the area of information technology, timesharing bureaux were common in the 1960s and 1970s and facilities management became common later on (Grover et al 1994). Software development has perhaps been the most widespread case of subcontracting for IT by all types and sizes of companies (Rands 1992). There are significant analogies between outsourcing various business activities and IT but differences are also substantial. Willcocks and Lacity (1995) and Alpar and Saharia (1995) have attempted a comparison. Among the differences they quote, three are highlighted. First, the elusive nature of information as a fundamental part of information systems. Second, the rapidly changing nature of information technologies and their characteristics. Third, the ubiquitous penetration of IT into all areas of business, making it difficult to draw a clear boundary around it. In general, inasmuch as outsourcing car parts is different to outsourcing catering services, IT outsourcing presents its own idiosyncrasies. Consequently, this thesis focuses on outsourcing IT services and the peculiarities

---

[7] *Computer Weekly*, 16 May 1996, p. 1.

[8] *Computer Weekly*, 5 September 1996, pp. 1, 29.

involved, while many of the discussions and propositions are expected to be more widely applicable.

The IT outsourcing market is still growing at an increasing rate, particularly after the well known case of Kodak (Loh and Venkatraman 1992a). According to International Data Corporation in 1991, the IT outsourcing market was expanding at a rate of 20% per annum and it was expected to reach £1013 million by 1995[9]. In 1992 The Yankee Group estimated that total revenues for all types of outsourcing totalled $33 billion and was expected to reach $49 billion in 1994[10]. The US IT outsourcing market was estimated at $22.8 billion in 1988, $29 billion in 1990 and was expected to reach $49.5 billion by 1994[11].

Hurst and Hanessian (1995) report that external providers command a substantial share of total spending on applications: 19% world-wide and 30% in the US, excluding PC applications. In 1995 the US market for IT outsourcing was the biggest, estimated to exceed $18.2 billion (Patane and Jurison, 1994). According to INPUT[12], the total IT outsourcing market in Europe was estimated at $4600 million in 1994 and at $5750 million in 1995. It is projected to reach $14300 million by the turn of the century. The 1995 figure comprises of applications operations (47%), platform operations (21%), desktop services (11%), network management (11%), business operations (5%) and applications management (5%).

Despite the well publicised successes, IT outsourcing has not been a pleasant experience for all the companies that have tried it. Lacity et al (1995) report that of 14 decisions to outsource the bulk of IT, senior managers declared 3 failures, 9 others seemed to be at risk of failing and only 2 decisions, involving data centres (a

---

[9] "The impact of facilities management on the IT market", International Data Corporation, London, 1991.

[10] *Network World*, February 1992, pp. 1, 31-36.

[11] *ComputerWorld*, 25 June 1990, pp. 122.

[12] *Information Systems Outsourcing Market Europe 1995-2000*, INPUT, London, 1995.

fairly standard service), were successful. Of the 15 decisions to keep most IT services in-house, 5 failed to produce the anticipated cost reductions or service improvements. The other 10 were successful from the senior management perspective but users felt that service levels had deteriorated.

As a response to the problems of total insourcing and total outsourcing, organisations now seem to be taking a middle road. In 1996 Dataquest published evidence showing that while the total number of IT outsourcing contracts continues to rise, the average value of contracts started to decrease in 1995. This indicates that total outsourcing contracts are proving less popular and companies are tending to favour more selective 'out-tasking'. This appears to be a more successful approach as suggested by the study by Lacity et al (1995). Of the 32 selective outsourcing decisions they studied, 20 met top management's objectives and 3 were complete failures. At the time of the study it was too soon to determine the outcomes of the remaining 9 decisions.

It becomes apparent that over time the practice of outsourcing has been continually changing. Sporadic contracts for software development and facilities management, gave way to total outsourcing 'mega-deals'; strategies ranged from selective 'out-tasking' to subcontracting the management of the whole IT supply chain. One way of capturing this moving target is to recast the research question in a more relevant way. It appears that the question as to whether companies should outsource their IT or not is too limiting and fails to give rise to adequate explanations of observed practices. The purpose of this chapter is to trace the progress made so far by academic IS research in studying the outsourcing phenomenon and generate some insight as to how should the research question be formed.

The following section attempts to consolidate the various proposed definitions of outsourcing in a critical manner. Section 3 reviews the drivers for outsourcing and the expectations companies have from it while also examining the actual benefits achieved or risks faced. Next, a number of decision making frameworks are

reviewed in an attempt to highlight the strengths and weaknesses of each with a view to demonstrating the evolving understanding of IT outsourcing within the IS community. Subsequently, the approaches taken and the advances made, in terms of applying mainstream economic theories to IT outsourcing, are examined. Section 6 looks at the prevailing theories on the structure and governance of the IT function and draws implications for our understanding of outsourcing. The chapter concludes by recapitulating the main theoretical advances so far and by setting the target of this thesis.

2. Towards a definition of outsourcing

A range of new 'buzzwords' have been introduced with the expansion of contracts for IT services. When 80% or more of an organisation's IT budget is outsourced, it tends to be called total outsourcing (Willcocks and Fitzgerald 1994a). However, outsourcing, insourcing, rightsourcing, smartsourcing (Willcocks and Fitzgerald 1994c), transitional outsourcing (Winder 1994) and out-tasking are some of these newly introduced terms, used to describe a variety of forms of IT contracting. They tend to be over-used and abused resulting in the loss of any specific meaning they might have had. These terms are employed on the basis of a generic, common sense understanding and in the absence of more exact terminology they are also used in research contexts where more precision is normally required.

In the broader business context, the Director General of the International Federation of Purchasing and Materials Management defines outsourcing as '*finding new suppliers and new ways to secure the delivery of raw materials, goods, components and services. [It means that you] use the knowledge, experience and creativity of new suppliers which you did not use previously*' (Rothery and Robertson 1995, p.4).

Besides such practitioner conceptions of outsourcing, academic researchers have attempted to provide more accurate definitions. However, even these attempts are

often constrained by the purposes of the research project from which they emanate. For example, some definitions limit the types of IT services they encompass.

In some instances, definitions of outsourcing are limited to particular activities or services. Rosenberg and Macaulay (1993) distinguish between physical (e.g. data centres, desktops, networks) and logical (payroll, billing, cash management) outsourcing, thus emphasising a technical and a functional perspective. Consulting, after sale services, and leased lines are excluded by Cheon et al (1995) and Grover et al (1996) who adopt a very broad definition of IS outsourcing as turning over part of all IS functions to an external provider.

Another narrow definition is given by Lacity and Hirschheim (1994). They focus on moving IT activities from in-house to external provision, thus excluding the use of contractors for new systems or services. In a similar effort, driven by a given research project, Heinzl (1993) distinguishes 'internal IS outsourcing' as the assignment of IS service provision to a subsidiary (a spun-off IS department or an acquired IT company, closely controlled through "*capital and personal ties*") from 'external outsourcing' (i.e. long-term service contracts to 'independent' suppliers).

A much more focused approach is followed by Alpar and Saharia (1995) who define IS outsourcing as an intermediate to long-term arrangement between an outsourcing firm and one or more independent vendors who are contracted to provide the firm repeatedly with various information systems services throughout the life of the contract. Resource transfer is not part of their definition which implicitly also excludes project-based contracts in favour of time-based agreements.

Often, a general description of the main apparent elements of an outsourcing deal is given in the definition. For example, Loh and Venkatraman (1992b) define outsourcing as the significant contribution by external vendors in the physical and/or human resources associated with the entire or specific components of the IT

infrastructure in the user organisation. Similarly, for Clark et al (1995) outsourcing is the delegation by contractual arrangement, of all or any part of the technical resources, the human resources and the management responsibilities associated with providing IT services to an external vendor. Kern and Willcocks (1996) define IS outsourcing as a decision taken by an organisation to contract out or sell the organisation's IT assets, people and/or activities to a third party vendor who in return provides the services, for a certain time period and monetary fee.

The above definitions tend to describe the relevant decisions or actions taken by companies in a general way and do not specify what distinguishes outsourcing from other forms of contractual agreements between IT suppliers and client organisations. Moreover, no systematic comparative studies have been found in the literature examining the differences between what is currently understood by outsourcing and other, more traditional (before the term outsourcing became common), forms of IT contracting.

Nevertheless, some important contributions have been made by the introduction of the notions of management, control and ownership in definitions of outsourcing. According to one such definition, outsourcing is "*the commissioning of third-party management of IT assets, people and/or activities to required result. This can and often does involve a degree of transfer of assets and staff to the third-party organisation*" (Willcocks and Fitzgerald 1994a, 1993a; Fitzgerald and Willcocks 1994a). Within the same logic, Feeny et al (1993) distinguish insourcing as the contracts according to which the market provides resources to be deployed under the buyer's management and control. In a similar but more formal way, Loh and Venkatraman (1992a) define outsourcing as the transfer of property or decision rights in varying degrees over the IT infrastructure by a user organisation to an external organisation such as a technology vendor or a systems integrator.

One common thread is readily apparent among the definitions reviewed above. IT outsourcing is a contract for the provision of IT services between two firms, namely

the user or client and the supplier or vendor. This immediately excludes at least one important contracting mode, that of fixed term (non-employment) contracts between a firm and an individual programmer, technician or consultant. This broad definition of outsourcing which is more formally specified by Loh and Venkatraman (1992a), seems to coincide with the common sense understanding of the term which tends not to exclude any forms of subcontracting, although the 'purest' and ultimate form of outsourcing seems to be total outsourcing.

Within this broad spectrum of contracting practices, certain defining parameters prevail. These are the notions of management, control, ownership, decision rights, contract characteristics, duration and number of services involved. These elements will be examined in later parts of the thesis. In the absence of a precise definition of outsourcing, the term is being used in this thesis in a generic manner unless otherwise specified. Whenever needed, reference is made to large 'total' outsourcing contracts, as defined above, or to other specific forms of contracting.

3. Drivers, expectations, benefits and risks

The study of the benefits and costs of outsourcing has attracted a lot of attention from IS academics. It is often observed that purported benefits of outsourcing in initially attractive deals turn out to become risks or symptoms of failure (Lacity and Hirschheim, 1993a). The typical example is that of heralded cost reductions. Some companies fail to see cost reductions and they often experience cost increases. Another example is the claim for flexibility. Apart from the issue of technical flexibility discussed previously, commercial flexibility is also at risk as outsourcing decisions, particularly large ones involving all IT operations, are difficult and costly to reverse or even to switch to alternative suppliers (Willcocks and Fitzgerald, 1994c).

In such a relationship where both parties are tied to each other, the client is subject to supplier opportunism which is usually manifested in subtle ways rather than in terms of major contract breaches. Since IT contracts are inherently incomplete,

vendors are particularly keen to designate a particular user demand as work that is out-of-the-scope of the contract that has to be charged separately. This is a source of increased costs and subsequent friction in the buyer-supplier relationship. A typical example of the conditions that emerge in large outsourcing arrangements is described in the following quotation by the Director of Business Services of the Inland Revenue in the UK, two years into the contract (Macintyre, 1996):

> *"... in general things are going well ... on the days when it gets a bit rough, and it does get a bit rough, we have to constantly remind ourselves that we are heading towards the same goals."*

Some commentators in the field suggest that outsourcing does not solve any problems or troubles that a company may face with the management of IT. On the contrary, it is likely to exacerbate any such problems because a supplier has little or no incentive to improve a problematic condition as long as it is profitable for them. Typically, some sort of baseline measurement is employed to determine service level requirements. A contract specifying service levels set on the basis of a problematic operation, only ensures that the operation remains problematic. Yet, many companies appear to resort to outsourcing as a means of correcting their existing IT operations, or getting rid of the problems they face. Moreover, client organisations sometimes lack the necessary expertise to manage such contracts and fall prey to the vendor's experience with a number of clients. The importance of management involvement is highlighted by the research of Saarinen and Saaksjarvi (1993). Nonetheless, from a survey of Fortune 500 companies, Loh and Vankatraman (1995) found that IT outsourcing is, overall, being successfully exploited in the sense that it is associated with improved firm performance, despite the large number of threats and risk that have been identified (Peak and Windson 1993, Peak 1994, Earl 1996).

Following the research by Gupta and Gupta (1992), Ketler and Walstrom (1993), Clark and Zmud (1993), Lacity et al (1994, 1995), Venkatraman and Loh (1994), Khosrowpour (1994), Fitzgerald (1994), Clark et al (1995) and Grover et al

(1996), the drivers of, or expected benefits from outsourcing can be consolidated under four categories: financial and economic, business (including strategy and organisation), technological and political. In the following sections, these are considered in turn.

3.1. Financial and economic expectations

The financial costs and benefits of IT outsourcing are widely contested (McLellan et al, 1995). Perhaps the most commonly quoted reason for outsourcing is cost reduction. These cost reductions are said to be based on economies of scale and scope that specialised suppliers can achieve. The existence of such economies is, however, questionable. There are relatively few studies estimating scale economies (e.g. Barron 1992, West 1994) and they are limited to the performance of mainframes (e.g. Mendelson 1987, Minoli 1995) and to software development and maintenance (e.g. Banker and Kemerer 1989, Banker and Slaughter, 1994). However, mainframe operation and software development are only parts of a multitude of IT services being contracted out. Moreover, from those studies, it appears that mainframes display constant or diminishing returns to scale and that the efficient size of mainframe capacity in MIPS is fairly low and easily attainable even by medium-sized companies. On the other hand, there seem to be substantial economies of scale in software development and maintenance, primarily due to learning effects and the use of formal tools and techniques.

Large corporations can often achieve a large scale on their own. They can thus provide the services they need as efficiently as the market. Telecommunications and wide area networks have been such cases before the deregulation of the respective markets. Large companies had their own, proprietary telecommunications infrastructure as prices were set too high by state-owned monopolies. Nonetheless, such multinationals were the ones that first initiated large outsourcing contracts, when the telecommunications markets started becoming competitive. Recent developments in telecommunications technologies also spurred further growth in

the market. In this respect, empirical evidence seems to support Adam Smith's proposition that the division of labour is limited by the size of the market (Stigler 1968), while showing that the presence of scale economies on its own is not a sufficient determinant of outsourcing.

Apart from scale economies, resource utilisation appears to be a fairly plausible reason for outsourcing, particularly for those IT services which are characterised by significant demand variations over time. For example, demand for mainframe access and processing power may be higher during the end of the financial quarter or year but significantly lower the rest of the time. Similarly, systems development skills are required in large numbers only for the duration of the project. By outsourcing a poorly utilised resource companies can limit their cost to the amount of the resource they actually use. This tactic presupposes that the supplier can balance the resource demands from its various customers. If not, cost reductions cannot be achieved. Thomson Holidays decided to outsource mainframe operations (but not the development and maintenance of the respective software) partly because they make little use of mainframe capacity during slack periods of the year but need considerable capacity to cope with surges in demand for holidays[13]. According to the same news article, IBM (one of the vendors invited to tender) found that Thomson's busy periods coincide with slumps in mainframe demand from other customers.

It appears that some of the expected cost reductions are not based on hard evidence but rather on the promises of suppliers who bid aggressively for IT service contracts, in anticipation of a long term contract which would subsequently bring them in superior negotiating position (Whang 1995), The supply of hardware and software has become a low-margin business for vendors (IDS and WBS 1994). Bundling the provision of hardware, software and support services, which have

---

[13] Collins, T (1997), "Thomson plans a holiday from running mainframes", *Computer Weekly*, 13 February 1997, p. 4.

varying profit margins, into large outsourcing contracts appears to be a more profitable strategy for them as it allows them to offer a competitive overall price while ensuring acceptable profit. Amongst the services currently offered by outsourcers, some (e.g. desktop support) are considered more lucrative than others (e.g. data centres). The marketing strategies of suppliers may therefore play a significant role in the decision of potential clients to outsource.

The anticipated cost reductions from outsourcing IT rarely materialise. Either because the vendor's original bid was unrealistic or because suppliers do not enjoy inherent cost advantages, client organisations have faced less than expected savings or even cost increases (Lacity and Willcocks 1995, Willcocks and Lacity 1995). Failed expectations are not the only sources of increased costs. Incomplete contract specification is another reason. Quite often companies find themselves charged separately for work not specified in the original contract. For these additional, incremental costs, suppliers take the opportunity to charge relatively high mark-ups (Willcocks and Fitzgerald 1994c).

A purely financial reason for outsourcing is that companies typically receive a substantial amount of cash in return for selling their IT assets to the supplier. This has several positive side-effects on their accounting books (Clark and Zmud 1993, Grover et al 1996). First, the net cash infusion can be used to finance other, perhaps problematic, areas of the business. Second, the company reduces its capital asset base. Third, a large number of employees are eliminated from the payroll as they are usually transferred to the supplier.

Other financial benefits reportedly expected from outsourcing include the conversion from fixed to variable costs, greater predictability of IT expenses and better accountability and control over costs. Sometimes, these expectations also fail to materialise. Although capital costs are reduced, costs rarely become truly variable as contracts typically specify fixed fees. Expenses are also not entirely predictable, mainly due to additional work, not specified in the contract (Willcocks

and Fitzgerald, 1994c). The formality of accountability and control as opposed to informal and ad hoc procedures in traditional IT departments is, however, a more universally admitted benefit of outsourcing.

3.2. Strategic and organisational benefits and risks

Outsourcing is considered to have the potential to provide a basis for organisational alignment, mainly in terms of incentives and in terms of business and IT strategy formulation (Kambil and Turner, 1993). The concentration on core business activities has been a major theme in management since the late 1980s. In the context of this line of thought, proponents of outsourcing have argued that (non-IT) companies are not in the business of building and maintaining systems and should therefore buy-in all related services without having to maintain the in-house capability. There are two fundamental problems with this argument. First, it assumes that there is a market for those services. While some question the maturity of this market, it appears that such a market exists today (at least for some IT services), though it may have not existed ten years ago. Moreover, many managers and academics alike, recognise that for certain services, specific to a company, a market may still not exist. Second, it may be the case that at least some parts of IT services are so closely intertwined with other business activities that it is difficult to separate them.

A similar point is made by those who argue that it is not the technology as such that creates competitive advantage, it is the information and the way it is used that makes the difference (e.g. Hopper 1990, Vowler, 1996b). The information within an information system would possibly embody company-specific knowledge and expertise which can lead to comparative advantage. It is not clear however, that this information can always be separated from the physical systems that handle it and the supporting services. An argument could be made that the software and hardware used may also embody some intelligence and expertise or that, at least, they create some of the desirable attributes of information (such as speed, accuracy etc.). Having said that, an increasing number of organisations are effectively

separating the business benefit from the technology. For example, Lufthansa is procuring most of its IT services from its IT subsidiary which has been spun-off. They are now trying to market Sysnet Netsched, their proprietary network management system to their competitors, while explicitly acknowledging that network management is the most essential core operation of any airline company[14]. A few years ago, American Airlines was making a similar statement regarding the handling of 'core' or 'strategic' information systems when it outsourced and opened to its competitors the famous SABRE system, making it their largest source of revenue (Hopper, 1990).

Outsourcing is also employed as a means of facilitating the transition from one technological architecture to another. The typical example is that of outsourcing mainframe operations while developing a client-server infrastructure (Minoli 1995). Such transitions may be of a major and strategic nature or of a more tactical character.

Expected improvements in the productivity of the IT group appears to be an important reason for outsourcing. The applications and maintenance backlog has been a major problem for many organisations, particularly during the 1970s and 1980s (Nolan 1979). Outsourcing is typically expected to help improve productivity by reallocating the effort and attention of the IT department and by relying on the alleged productivity of professional service teams and companies. The initiation of an outsourcing exercise on this basis by management is likely to induce a competitive bid from the internal IT group, wishing to prove its own efficiency (Lacity and Hirschheim, 1993b). Indeed, the threat of outsourcing can, at least in theory, be an effective incentive for the in-house group (Wang and Barron 1995, Lacity et al 1996).

---

[14] "Refuelling Lufthansa", *Information Strategy*, Vol. 2, No. 3, April 1997, pp. 39-41.

Another organisational reason for initiating outsourcing exercises is the market testing of internal IT operations. Some companies issue invitations to tender without a clear intention to outsource. Instead, they seek to assess their internal operations as well as to collect possible innovative ideas from vendor bids. In response, suppliers often refuse to bid or withdraw their proposals, as mentioned on page 3.

3.3. *Technical and Technological reasons for outsourcing*

Acquiring access to technical expertise and to leading edge technologies is claimed to be a good reason for outsourcing. Assuming that vendors can keep pace with technological developments and that they maintain highly skilled employees, client organisations can benefit from handing their IT systems and services over to them. This argument is partly true, especially with emerging technologies.

Reducing technological risk and improving flexibility and service levels are oft-quoted reasons for outsourcing (Lacity and Hirschheim, 1993b). According to this argument, vendors enjoy inherent technological flexibility and technical superiority so that a client can be better protected from technological obsolescence or poor service performance (Duncan, 1995). The truth appears to be that vendors can be as locked-into a given set of technologies and standards as any client organisation. Switching to alternative technologies is a costly and risky activity. Moreover, large and complex outsourcing contracts rarely provide the flexibility to make such transitions. Companies may become locked into the supplier's technology and then observe that other companies adopt alternative, lower cost and higher quality systems.

The management of diverse resources is a more plausible reason for outsourcing. Information technologies keep expanding so that organisations have to manage an increasing variety of technologies, standards and systems. New generations of technologies do not always replace earlier generations but often complement them. This seems to be the case with mainframes, personal computers and network computers, for example. The task of managing the increasing variety of human and

physical resources needed to handle an increasing number of technologies becomes overwhelming. Outsourcing some or all IT services is a means of containing the expansion in management and technical human resources.

3.4. Political antecedents to outsourcing

Outsourcing is often initiated as a result of internal political manoeuvring between managerial ranks. The IT department often finds itself in the middle of conflicting demands from users and general management. Users constantly request more systems and services of a higher quality and technological standard, more promptly. General management is putting pressure for cost reduction as the IT budget keeps increasing while any benefits cannot always be assigned specific monetary values or improvements in business performance. The IT department is in the middle of these conflicting demands and, as a result, is perceived by both user and senior management as a troublesome function that they do not understand (Lacity and Hirschheim, 1993b)

General or financial managers call in outsourcers in order to control costs and/or to get rid of the problem. CEOs can be tempted by outsourcers promising substantial cost reductions and cash transfers in return for taking over IT operations. They may also be tempted by the hyperbole in the trade press and the 'bandwagon' (Lacity and Hirschheim, 1993a) effect created by early, highly publicised outsourcing deals (Loh and Venkatraman, 1992a).

In this context of conflicting demands, IT directors or managers also have various incentives for initiating outsourcing. Proving their efficiency and effectiveness is one reason. Justifying further resource demands is another (the opinion of an external consultant may be perceived as more credible expert advice). Raising their profile and status within the organisation is a third.

Lacity (1992) has extensively studied the political interactions affecting IT outsourcing decisions and has argued that they often bear a heavier weight on the

final decision as compared to 'rational' economic criteria. An outsourcing decision which is solely a result of political manoeuvring is likely to fail to yield acceptable service levels and costs. It is an important decision with substantial risks and as such requires significant management attention and skill.

4. The evolving conceptualisation of IT outsourcing

A large amount of research effort in the IS field has been devoted to the study of the management decision to outsource IT services. Various theoretical propositions and empirical accounts have been submitted, addressing the decision making process, the criteria employed, and/or the types of outsourcing arrangements available. It has become evident that the organisational processes leading up to outsourcing decisions are very complex, involving a number of stakeholders with different agendas, trying to find a balance between costs, performance, flexibility and long term viability (Macey 1993, Willcocks et al 1995, 1996).

Several early accounts of IT outsourcing have emphasised the partnership dimension in such deals. In recognition of the complexity, uncertainty and idiosyncrasy of the provision of IT services, management consultants, the trade press and academics have discussed the need for mutual trust, a collaborative culture and other similar qualities assigned to a fairly broad notion of partnership (Willcocks and Choi 1995, Klepper 1995). Partnership has also been associated with strategic alliances, which, in turn, also received substantial attention (e.g. McFarlan and Nolan 1995, Willcocks 1994). It has become clear, however, that a partnership attitude, although valuable, is not a substitute for a good contract (Panday 1995). Moreover, it can be argued that what is described as partnership in businesses, refers, at least to some extent, to institutional structures such as joint decision making and arbitration, which are typical of relational contracts[15].

---

[15] More on the governance structures of relational contract in section 5 and chapter 2.

Often, a bipolar distinction is made, between internal IT provision and outsourcing. Once these are recognised as the only two options available, the question becomes one of how to select between the two. Responses to this question range from relatively simple checklists, such as the questions that Benko (1992) and Jones (1994) ask, to more elaborate frameworks. For example, Grover and Teng (1993) propose a two-level analysis framework which considers characteristics of the systems (such as maturity, significance and existing capability) on one level and the impacts (costs and benefits) of the decision on the other. From a slightly different viewpoint, Cronk and Sharp (1995) argue that whether a system (or service) should be outsourced or not depends on whether it is a *unit of competitive advantage*, a *value added* or an *essential support* system. These approaches emphasise those characteristics of the system or service that make it more or less important for the organisation.

More technically 'hard' approaches have also been developed. Clark and Zmud (1993) tested a large causal model of the outsourcing decision structure, taking into account about 55 variables. Putrus (1994) and Alpar and Saharia (1995) have recommended the use of the Analytic Hierarchy Process as a means of resolving the various trade-offs between using internal and external groups for IT provision.

While the above studies assume the existence of only two possibilities (internal or external organisation of IT provision) many researchers in the field, have recognised the existence of a variety of possible outsourcing arrangements which widen the options available to an organisation and can make a difference to the success of IT provision. A descriptive framework for classifying IT outsourcing possibilities has been proposed by De Loof (1995). It characterises alternative options for organising the provision of IT according to elements of the IS function, the provider, the relationship between supplier and client and the allocation of ownership and control over IS components (systems and services).

Several classifications claim to be more than descriptive and attempt to prescribe forms of contracting on the basis of some selected criteria. Such a classification of IT outsourcing arrangements has been proposed by Ang (1993). Specifically, six types of providers or modes of organisation are proposed, namely, in-house, parent company, other company in same industry, service bureau, joint venture and facilities management. The determinants of this classification are the spatial location of the service, the legal ownership of assets, the legal status of human resources and the behavioural control (managerial discretion) exercised by the client. The focus here is not on system characteristics, but on organisational and contractual issues. Apte and Mason (1993) further emphasise the geographical dispersion of IT service provision in large corporations, coining the term 'disaggregation'. Outsourcing often involves relocating services such as data centres and help desks, perhaps on a global scale, but this is not necessarily the most significant parameter in the outsourcing decisions.

Fitzgerald and Willcocks (1994b) combine the uncertainty involved in an outsourcing deal with the degree of contractual definition to recommend effective contract structures. In particular they recommend a fixed fee, tightly specified contract under low uncertainty, or a loosely defined, risk/reward sharing contract under high uncertainty.

At a general level, Nam et al (1995) take into account the strategic impact of IS and the degree of substitution of supplier for in-house services to propose a two-dimensional grid of types of outsourcing relationships, namely, support, reliance, alignment and alliance. In a more detailed approach, Willcocks and Fitzgerald (1993b) and Willcocks et al (1994) use business uncertainty and system interconnectedness, clients' experience with technology and service specificity, as criteria guiding the selection of an appropriate approach to the organisation of IT provision. Both approaches emphasise that the characteristics of the system, the capability of client and supplier and the conditions of the commercial environment co-determine the outsourcing decision.

One problem often encountered in such exercises, is the diversity of systems and services managed by organisations. Producing a contingency framework covering all these systems and services is, therefore, problematic. As a result, some studies focus on narrower sets of services. Concentrating on applications development, Buck-Lew (1992) distinguishes among three modes of organising its provision, namely, pure and hybrid outsourcing and in-house development. For the same types of services, Saarinen and Vepsalainen (1994) propose a framework according to which applications of different levels of uncertainty and specificity are matched to three procurement options, namely, off the shelf software purchases, contracting out and in-house development. Similarly, Saaksjarvi (1993) matches four IS development roles (design, implementation, operations and support) to four sourcing alternatives (centralised, decentralised, partnership and service market) depending on three business management criteria (flexibility, innovativeness and efficiency) and three IS management criteria (joint planning, resource allocation and performance evaluation). Interestingly, Iivari and Ervasti (1993) found that systems development is more effective when carried out in-house, depending on the complexity of systems.

In a similar effort, Klepper (1993) draws on Williamson's (1985) contingency framework for classifying vertical governance structures, to derive implications in the context of organising IT systems and services. Klepper notes that the framework as such provides unambiguous recommendations only for two, rather extreme cases, namely that a classical contract is appropriate when both uncertainty and asset specificity are low and that a relational contract is needed when uncertainty, frequency and specificity are high. For intermediate positions, Klepper suggests that that decision should depend on individual cases. Most instances, however, would be expected to lie in the middle ground between the two unambiguous extremes. Moreover, the categories of contract supplied in Williamson's framework may be too broad to form the basis of usable recommendations. In particular, relational contracts comprise a very broad category.

Loh (1994) proposes a more elaborate model of IT governance incorporating transaction costs, bargaining costs, agency costs, influence costs, production costs, decision information costs and management costs. Equivalent work by Aubert et al (1994) includes asset specificity, measurement problems and origin of the most important investment (business or technical skills).

Essentially, all these studies attempt to find a balance between comprehensiveness with a very large number of determinant variables (e.g. Clark and Zmud 1993) and easy conceptualisation and implementation of the decision making framework. Moreover, they try to classify the organisational and contractual possibilities available to a limited menu of options. It appears that this quest remains incomplete, since both the determinants and the options available, seem to be too complicated to be captured in a simple, formalistic framework.

On a different research direction, Loh and Venkatraman (1995) discuss the move from a dichotomous insourcing versus outsourcing decision problem to the concept of 'rightsourcing'. Following Henderson (1990) they recommend a portfolio approach to IT supply relationships where emphasis should be placed on 'co-sourcing'. Their notions of rightsourcing and co-sourcing refer to a loosely defined strategy of engaging in cooperative supply relationships with various suppliers and under different contracting modes, depending on the characteristics of the systems and services being outsourced. In particular, they have ranked a number of IT services according to the degree of internalisation of physical and human resources by the provider, the degree of specificity of the business domain and the extent to which a contract is project or period based (Loh and Venkatraman, 1992b).

The theoretical understanding of the IT outsourcing phenomenon appears to have evolved with developments in the marketplace (Wassenar and Thiadens, 1993). While the outsourcing market continues to grow, both in value and in number of deals, it appears that the average size of each deal decreases (Takac, 1994).

Venkatraman and Loh (1994) argue that over time, IT provision arrangements are converging to an equilibrium in the region between 'total' insourcing and 'total' outsourcing which have prevailed in the recent past. They characterise this position as 'relationship portfolio' where IT competence is shared between the client organisation and a number of suppliers. Despite its abstract form that provides little guidance to researchers and practitioners, this framework succinctly summarises the departure from the bipolar dilemma between total insourcing and total outsourcing.

The notion of selective outsourcing as described by Lacity et al (1995) is also based on the premise that 'total' solutions -where the majority of IT services is provided by a single provider, whether it is the internal IT department or an external supplier- are not efficacious and that small parts of all IT activity should be awarded on multiple providers. The fundamental contribution of the notion of selective outsourcing to our understanding of the possibilities available for the organisation of IT provision remains significant, despite the unrealistic claim that each service can be awarded to more than one supplier and the limited extent to which IT services can be split into very small pieces.

Ang (1994) took these ideas one step further and studied the degree of managerial control by the client and the supplier for seven IS activities under six different modes of contracting. This analysis clearly shows the difference that both the nature of the service as well and the contractual mode make on the allocation of decision and control rights for IT. The message coming out of such studies is that different IT services pose varying requirements for the organisation of their supply while different contractual arrangements afford varying capabilities in terms of control and decision rights. This argument is wholly consistent with the basic premise of transaction cost economics (Williamson, 1985) which is more fully reviewed in chapter 2.

These developments raise a whole new set of issues regarding the organisation and management of the remaining IT function within the client company. Loh and Venkatraman (1995) argue that the IT function should not be seen as the 'glasshouse' any more but as the 'clearing house' for IT services. This will involve a substantial shift from technical skills to contract management and relationship building skills. From a case study analysis of 30 residual IT functions, Willcocks and Fitzgerald (1994d) have identified seven major roles or tasks retained or bought in the firm after outsourcing, namely, IT strategy, business development, systems integration, informed buyer, contract intermediary, contract monitoring and exploitation of contract relations.

In addition to academics, consultants also share similar views for the reshaped IT function after outsourcing. Aylott (1995) envisages a tripartite relationship between the IT function, the users and the supplier where each pair is involved in a different kind of exchange. However, in practice such relationships are more complex as more than one suppliers are typically involved, interacting with multiple IT units and user groups in large client organisations. In its guidelines for successful outsourcing, KPMG[16] classifies the tasks of the client organisation's IT group into management issues (related to planning, negotiations and contract management), human resource issues (related to in-house, previously in-house and the supplier's staff), service/business issues (related to service level agreements and service delivery) and communication/ understanding issues (related to management coordination with the supplier).

Acquiring, developing and maintaining the skills required by the IT function in an environment where the majority of services have been outsourced, is a difficult management task in itself. Inasmuch as staff transfer from the client to the vendor serves the transfer of client-specific knowledge, the transfer of selected staff from the vendor to the client, serves a similar purpose, mainly in terms of contract

---

[16] "Best practice guidelines for outsourcing", KPMG IMPACT Programme, London: HMSO, 1995.

management skills (Willcocks and Fitzgerald, 1994c). It is not uncommon for vendor account managers to be employed by the client, as a result of the knowledge and experience they acquire over the course of the contract.

5. Economics in IT outsourcing research

The issues surrounding the IT outsourcing phenomenon naturally impinge on the question of 'make or buy' which has long been a central theme of several strands of economics. Among the first to point out the relevance of economic theory in IS research, were Kriebel and Moore (1981). They emphasised the vast amount of published work in economics faced by IS researchers and the problem of capturing and transferring this knowledge into the field of information systems. Bakos and Kemerer (1992) re-addressed this association between information systems and economics and highlighted outsourcing in their economics-in-IS research agenda. Many IS researchers have drawn on theoretical insights, methods and techniques found in economics as a basis of their own work on IT outsourcing.

Much of the groundwork required for economic analyses in IS research has been reviewed by West (1994). He discusses the effect of various sources of production costs (such as labour specialisation, employment of capital, public goods etc.) on seven functional components of information systems (such as hardware, software, telecommunications etc.). Moreover he points out that decisions regarding the costs of one component will affect the costs of other components and consequently the costs of the whole system, thus highlighting the economic inter-dependency of IT services.

In the context of IT outsourcing, West (1994) argues that it is characterised by high transaction costs and that it is only rational if it provides production cost savings which at least offset the increase in transaction costs. This proposition appears to be partial in that it does not explicitly take account of the importance of contract structure in shaping the realised transaction costs. It also appears to be based on a

questionable hidden assumption that production and transaction costs are substitutes. It can be argued that either production or transaction costs may render certain governance options altogether infeasible or at least highly inefficient, although the increase in one can be exceeded by the savings in the other.

Indeed, if transaction and production costs were substitutable, the same argument could be used for both insourcing and outsourcing: if the transaction cost savings of internal organisation exceed production costs then insourcing would be preferable. Lacity and Hirschheim (1995) have examined such cases where insourcing was chosen as the preferred option. However, internal organisation is known to have various inherent disadvantages as compared to market transactions, mainly incentive slack (Williamson 1995) and influence costs (Milgrom and Roberts 1992).

Transaction cost economics has been a popular reference theory among IS researchers studying IT outsourcing, presumably because it directly addresses issues of vertical integration and make-or-buy decisions. In this theory, asset specificity features as the main criterion. Alpar and Saharia (1995) quote Reeve (1990) in arguing that highly specific IT activities can be broken down into sub-functions, each of which can potentially have lower specificity. Of course, they admit, the coordination across sub-functions may in itself be characterised by high specificity and therefore may have to remain part of the in-house operations or it may not allow the separation of sub-functions. Indeed, management coordination turns out to be an important determinant of IT governance, as shown in subsequent chapters.

Software development may be considered such a sub-function. Following a similar line of argumentation and drawing on contract theory, Ang and Beath (1993) discuss in depth the relational characteristics of software development contracts. In particular they focus on hierarchical elements, typical of internal organisation, that are found in such contracts. Their analysis constitutes further testimony to the fact

that the important question is not so much whether to outsource or not but how to set up the respective governance structure, whether within the organisational hierarchy or across the firm's boundary. This has also been the fundamental conclusion of the study by Aubert et al (1996).

Often, transaction cost economics is criticised for a number of shortcomings. As a response, researchers may choose to couple insights from this theory with contributions from another in order either to compare the respective interpretations of IT outsourcing or to present a more balanced account. For example, Lacity and Hirschheim (1993b) demonstrate how the same evidence can be understood differently from a transaction cost perspective and from a political theory viewpoint. Political antecedents of outsourcing have been reviewed in section 3.4 above. In contrast, Jurison(1995) combined transaction cost principles with financial risk theory in an attempt to provide a more complete model of cost and risk differentials between the internal and external organisation of IT services. Cheon et al (1995) reviewed resource based theory, resource dependence theory, transaction cost economics and agency cost theory and synthesised them into a contingency model addressing IT outsourcing. From a slightly different angle, Teng et al (1995) explore the strategy-theoretic factors associated with IT outsourcing decisions by employing resource based theory, resource dependence theory and strategic orientation theories.

One of the most extensive critiques of the use of transaction cost theory in the context of IT outsourcing has been raised by Lacity and Willcocks (1995). They report 'anomalies', countering typical transaction cost theoretic predictions. They attempt to explain these anomalies by challenging some explicit or implicit assumptions of transaction cost economics, which seem to be profoundly violated in the case of IT services. Three of these violated assumptions are particularly worth mentioning. First, Lacity and Willcocks (1995) argue that for decision makers in the 61 outsourcing cases they studied, economic efficiency was not, on the whole, the primary criterion for their outsourcing decisions. Second, they found that

information is not always used to make rational decisions. Third, the transaction is not found to be an appropriate unit of analysis. This criticism is important but still open to debate, at least until a better theoretical alternative is brought forward.

Economic efficiency may not be the only driver of the outsourcing decision, as also discussed in previous sections. It is, however, reasonable to expect that whatever the primary driver, the costs and benefits must be part of the decision since economically inefficient operations are not viable in the longer run. The rationality of human decision making has also been widely debated (e.g. Simon et al 1992, Lawson 1997). That said, the position that human decision makers are generally irrational is even more open to criticism and requires elaborate qualification. The complication can be overcome (at least temporarily) by postulating that decision makers are in principle boundedly rational but try to maximise a set of subjective gains within a very broad agenda of economic, political, sentimental and other criteria. Thus, several apparently irrational decisions and actions can be attributed to subjective valuations of non-obvious alternatives. These valuations can still be internally (or subjectively) rational.

Lacity and Willcocks (1995) contend that the transaction is not the appropriate level of analysis, because it may be used to refer to different levels of aggregation in different cases. For example, some organisations outsource the whole IT function in one transaction whereas others outsource individual services selectively. The argument adopted in this thesis and developed in the next chapter is that whatever the level of aggregation in an IT supply transaction, it should be studied as an 'open box' with emphasis on the characteristics of its constituent parts and of the governance structure. Thus, the transaction as a unit of analysis is not a minimal datum but a rich environment for study.

From the realm of economics, several researchers have drawn on more mainstream modelling approaches, namely principal-agent and game theoretic models or, more generally, 'mechanism design'. The contracts for software development have been

studied more extensively. For example, Whang (1992) develops a game theoretic model of software contracting and suggests optimal incentive contracts, taking into consideration the existence of multiple stages in the process and the fact that costs and benefits are privately revealed to the developer and the user respectively, gradually over time. A damage clause is incorporated to allow for premature contract termination.

Some interesting results have been found by Richmond et al (1992) who have developed an incomplete contracting model to analyse information asymmetry, profit sharing rules and specific investments in software development outsourcing. Their main finding is that the net benefits to users from outsourcing are neither uniformly negative nor positive. Whether outsourcing is preferable to internal development or not depends on a number of issues. First, outsourcing can result in a higher value to the users since it provides incentives for specific investments which can lead to future cost reductions. Moreover, if the true cost and value of the project is common knowledge (between users and developers) then the incentives afforded by outsourcing are even stronger. However, when the costs and benefits are privately known to developers and users respectively, then each party has an incentive to misrepresent this information, and as a result users and developers may fail to reach certain otherwise beneficial development agreements. Internal development dominates outsourcing when either the developer's incentive to misrepresent privately known cost information is very strong and/or when there is a specific investment to be made by the users and it is comparatively more significant than any specific investments needed by the development group. If outsourcing is adopted, however, users should be prepared to give up some of the savings or benefits from the project, otherwise the developer's incentives will be detrimentally impaired.

These findings summarise some fundamental trade-offs involved in deciding between internal and external organisation of IT activities. Richmond and Seidman (1993) have also introduced the role of competition between suppliers for

undertaking sequential stages of the development process, in the study of these trade-offs. Although competition reduces prices, it also diffuses the incentive for value adding investments. Besides, an incumbent supplier in such cases acquires an advantageous position over rivals, both from its own as well as from the user's perspective.

Similar results are derived when competition is combined with 'learning by doing' and 'code reusability' effects (Whang 1995). In the presence of market competition, cost reductions due to these effects are not passed on to customers as each developer tries to underbid the other in order to win the first projects and build experience and reusable code for future projects. In so doing they are willing to accept smaller profits or even incur losses (low balling) in the beginning. Therefore, while development cost decreases over time, profit margin increases and price does not necessarily decrease. Users can reduce the prices quoted by potential suppliers by giving out more information about the project, thus reducing the supplier's cost uncertainty. Having said all that, it does not necessarily follow that eliminating competition should reduce the price paid for development.

Services other than software development have not been widely addressed in this stream of the literature, with few exceptions. The study by Wang and Barron (1995) is based on principal-agent models. They focus on 'processing power' services (e.g. data centres, mainframe capacity) where assets are non-specific, measurement is not a big problem, the market is competitive and queuing delays are pertinent. According to Williamson (1995), in the absence of transaction costs production should be allocated to the most efficient producer. Wang and Barron's (1995a) modelling and results are consistent with this principle. Their main finding is that central management should consider organising the IS department as a profit centre while exposing it to the threat of external competition in order to generate efficiency incentives and to simplify the outsourcing decision. Whenever the internal department doesn't break even in face of external competition, (some of) its activities should be outsourced. However, the same authors have shown that

a cost centre is superior to a profit centre in that it maximises the total value of the IT resource, despite the incentive slack (Wang and Barron 1995b). In practice, organisations do not appear to go through such an exercise before outsourcing their data centres. Even if it can be shown that internal departments cannot compete on price, this corroborates the finding that economic efficiency is not the only or necessarily the most important determinant of outsourcing. Moreover, assuming that transaction costs are entirely absent, is an exaggeration. There are specific investments associated with data centre operations, mainly in human assets, as the Gloucester County Council case shows (see page 2). The interdependence between data centre operations and other services such as network support, create further problems of coordination and management.

Further insights are generated by the study of Van Alstyne et al (1995) who have applied the theory of incomplete contracts on a number of issues related to database management. Their results are easily generalisable to other IT services. Two main principles are examined. First, if an organisation (more generally, agent) is indispensable to the functioning of a service, it should also own the respective assets, since ownership provides better incentives. Second, complementary services should be controlled together (see also Thong et al, 1994). There is a clear conflict in cases when there are more than one indispensable agents for the same asset or activity and when assets/activities are complementary and there are more than one indispensable agents. In such cases (which are expected to be fairly common in practice) there is no distribution of ownership (control) that induces first best investment. Moreover, there is no a priori precedence of complementarity or indispensability or vice versa. The second best balance is case specific and should give precedence to the principle that weighs most heavily in improving performance incentives while minimising the compromise on others.

The preceding discussion in this and previous sections reviews a number of criteria and trade-offs impinging on the outsourcing decision. Each of these theories has clear limitations but also highlights important aspects of the problem. The

arguments developed in this thesis draw on the lessons learned from the strengths and weaknesses of the theories reviewed above.

6. The governance of the IT function and outsourcing

Since the 1970s, the emphasis of debate on IT management has been on the question of centralisation versus decentralisation of IT resources and decision making. With the proliferation of end-user computing in the 1980s, another question surfaced, whether IT should be in the hands of specialists or users. These organisational designs for IT evolved as part of the data processing and MIS eras that marked that period (Earl 1989, Ward 1995). While there are efficiency and other reasons for centralising the provision of IT facilities, business users also require involvement in and control over some IT decisions. Still today, the relationship between users and IT is often fraught with problems (Ward and Peppard 1996).

With the growing recognition of the strategic importance of IT and the proliferation of strategic and competitive information systems (the strategic era) the size, the complexity and the budget for IT in large organisations increased dramatically. However, the disagreements over the allocation of resources, responsibilities and authority for IT did not seem to diminish. Although the question of whether to 'make or buy' IT services has always been an integral part of the structuring of the IT function within organisations, IT outsourcing has hit the headlines in the last ten years (1990s). As discussed in section 3 above, several senior IT or general managers saw total outsourcing arrangements as an easy way out of all these problems.

Having said that, these massive contracts opened a new range of contracting possibilities, which are markedly different from the traditional contracts for systems development or for temporary analysts/programmers. Despite the questionable managerial motives behind some of these exercises, long term service

contracts have their own merits, as discussed in previous sections. Initially the question was whether to outsource or to keep IT in house. However, one of the lessons we have learned from the debate on centralisation versus decentralisation is that polar extremes carry significant risks and that a balanced trade-off is more appropriate. By reviewing extant knowledge on the organisation and governance of the IT function, this section aims to draw lessons for the outsourcing decision.

The structure of the IT function ranges from the corporate centralised service, through the federal structure, to the completely decentralised (Hodgkinson 1996, Earl et al 1996). Admittedly, most large, complex organisations employ some form of federal structure, where internal bureaux, run as profit centres, or business ventures, run as semi-autonomous subsidiaries, are not uncommon. Research on (de)centralisation of IT highlights three aspects of the problem.

First, prior research has sought to examine what determines the structural alternatives adopted by organisations. The broader organisational structure and strategy as antecedents of centralised or decentralised IT, have attracted most interest and have yielded significant empirical support (Tavakolian 1989, Ahituv et al 1989, Allen and Boynton 1991). However, Brown and Magill (1994) and Brown (1997) observed that within the same multidivisional firm, the (de)centralisation of IT activities may not be uniform across business units. Organisational and strategic factors specific to each business unit and to the IT function, were found to be significant determinants of variations in organisational design among business units of the same firm. The theoretical concept underpinning these research efforts is the fit (alignment) between elements of the context (mainly organisational and strategic characteristics of the organisation, the business unit and the IT function) and the organisational design adopted for IT.

The second aspect of the (de)centralisation debate revolves around the categorisation of IT tasks, activities, decisions or services. This aspect, particularly

the significance and implication of the nature of IT activities, tends not to be directly addressed and is, therefore, under-researched. Although there is a lack of consensus on terminology and classification, most authors (e.g. Zmud et al 1986, Tavakolian 1989, Brown and Magill 1984 and 1998), broadly agree on a distinction between infrastructure activities (such as networks and communications) and user-oriented activities (such as systems development and maintenance). At the outset, for obvious reasons of efficiency and control, the former are better controlled as a central facility whereas the latter should be the responsibility of the users. This approach is usually taken for granted, with few variations. Von Simson (1990), who emphasises the distinction between responsibility for financial resources and responsibility for human resources, presents a different approach. His 'centrally decentralised' structure entails central control of human resources (recruitment, training, and career advancement) and decentralised control of financial resources and project priorities. The question of what is an appropriate classification of the things that the IT function does, remains unanswered.

The case study at Baxter Corporation (Main and Short 1989) demonstrates that without a closer look at the level of individual decisions, tasks and systems, generic classifications may not be sufficient either for research purposes or for decision making in practice. In order to avoid possible confounding effects of task-specific characteristics, some studies concentrate on one well defined task, such as systems development (e.g. Brown and Magill 1998) or data centre operation (e.g. Hufnagel and Birnberg 1989). In both studies, task-specific elements (e.g. user involvement in systems development and high fixed costs in mainframe operation) are evidently important but are excluded from the causal reasoning or models being developed. IT functions in large organisations today are involved in a number of processes (e.g. strategy, procurement, project management and security) and deploy a broad array of technologies (from 'legacy' systems to web-based intranets). A more systematic understanding of what makes an individual task or system more suitable for centralised or decentralised control is needed.

Finally, the third aspect of the debate is about what decision making and control mechanisms should be put in place to ensure the successful operation of whatever organisational design (centralised or decentralised) is adopted. Organisational structure is only one element of the governance of IT. Unless appropriate management processes complement it, the anticipated benefits may not be achieved and important strategic issues may be disregarded. Steering committees, functional and IT planning groups and chargeback procedures are some of the mechanisms commonly introduced (Earl 1989, Ward 1995). Moreover, as Brown and Magill (1998) found in their research, the determinants of organisational design for IT may be in conflict. In such cases they suggest that IT tasks and decisions may have to be further subdivided and the allocation of decision making may have to be further customised. Some of the complex intricacies involved in setting up such mechanisms are exposed in Taylor and Tucker's (1989) case study on centralising IT procurement decisions in a multidivisional bank. The study shows the range of issues (incentive, financial, political and others) that need to be addressed simultaneously and the variety of arrangements that can be introduced. Similarly, the case study by Hufnagel and Birnberg (1989) on chargeback for a centralised data centre, implicitly demonstrates that organisational design and management processes (all the elements of governance) need to be decided jointly since, to a certain extent, they determine each other.

This brief note on the issues of centralisation versus decentralisation and user versus specialist control serves to highlight that the questions we faced with the proliferation of large outsourcing contracts were not totally alien to previous knowledge and experience. The decision whether to centralise or decentralise IT resources can, to some extent, be paralleled with the decision whether to outsource or not.

Early work by Zmud (1984) put into perspective the organisational design of IT services. He noted the increasing complexity of the technological and market

environment and stressed that '*managing information resources within such an environment is considerably more difficult than the environments of the 1960s and 1970s, where an intimidating, relatively inflexible, but generally stable, technology base resulted in the centralised information systems authority maintaining a virtual monopoly as a provider of information service*'.

Since the early 1980s, this environment has become even more complex and unstable but Zmud's (1984) recommendation for organisation design is still valid, involving, '... *relationships between tasks and work groups such that work behaviours will be consistent with the performance criteria characterising that particular task environment*'. It is quite remarkable that, despite the age of this statement and its implicit focus on issues of centralisation versus decentralisation, it is still valid in the context of modern outsourcing. Namely, work groups, whether internal or external, need to be assigned to tasks (or services) by taking into account the particular characteristics of the latter and of the broader organisational structure and strategy.

The arguments developed by La Belle and Nyce (1987) also emphasise that dilemmas between polar extremes are false: "*the issue is ill served, we believe, by the label 'centralization vs. decentralization'. The game isn't tennis, it's a balancing act. [...] The question is how to judge what to centralize and decentralize*" (emphasis in original). The authors go on to analyse why and how different IT activities and services should be centralised/decentralised to a greater or lesser extent, according to their characteristics. This argument is equivalent to the notion of selective outsourcing (Lacity et al 1995). In tandem with others (e.g. Von Simson 1990), Earl et al (1996) note the continuous changes in the configuration of the IT function, consisting of minor reallocations of IT activities or adjustments in organisational structures and management processes. This discussion paints a picture similar to the notion of a portfolio of IT sourcing relationships (Venkatraman and Loh 1994) where the allocation of IT services to internal or external groups varies from one service to another and evolves over time.

Similarly valid for outsourcing is the framework by Feeny et al (1994) for assigning responsibilities for IT activities to users or specialists. Inasmuch as mature and low level (e.g. delivery and support) services should be the responsibility of specialists, they can also be outsourced. In contrast, less mature and higher level activities (e.g. strategy) should be the responsibility of senior user and IT management and should probably be kept in house.

In managing their IT resources, organisations have traditionally been separating and allocating IT activities to central divisions or to decentralised units, under the control of users or IT specialists, to be delivered by internal or external teams (Ward 1995, Cross et al 1997, Clark et al 1997). In other words, it has been a question of balancing the allocation of IT resources along three dimensions of organisational design. The proliferation of novel outsourcing practices in recent years has spawned renewed interest in the external dimension. Following the lessons learned in the study of (de)centralisation, there are three aspects of the external dimension that need consideration. First, the fact that the bipolar dilemma between the two extremes (insourcing vs. outsourcing) is false and that we need to examine the range of possibilities in the middle ground. Second and related to the first, the need to systematise the management processes, the decision making procedures and the other elements comprising the external governance of IT. Third, the need to explore those characteristics of IT systems and activities that render them more suitable for internal or external provision.

Extant literature has covered substantial ground, to varying degree in each of the three aspects. In these terms, the aim of this thesis is twofold: to *analyse* each of the three aspects in more detail and to *synthesise* the findings in one coherent and manageable framework.

7. Conclusions

A common thread permeating this chapter is that the way outsourcing relationships are established and structured is the fundamental determinant of success or failure. Whether to outsource or not is one question but a more important one is how to do it (e.g. Willcocks and Fitzgerald, 1993a). Kumar and Willcocks (1996) demonstrate examples of IT outsourcing attempts that went into trouble because of badly specified or non-functional contracts. Richmond et al (1992) have adopted the same position and, more specifically, argue that "*one has to determine the characteristics unique to each function in the information systems organisation and develop models that highlight these unique characteristics*". Once it is recognised that IT services vary both within and across organisations, different contractual and organisational structures ought to be devised to match their individual characteristics.

Moreover, Grover et al (1996) recognise that common classifications of IS outsourcing functional components (i.e. functional breakdown of IT activities) have limited granularity and that a closer and more detailed look at the characteristics of IT systems and services is warranted. This is similar to the principle of Van Alstyne et al (1995) that interdependent services should be identified and organised together. In the same line of thought, Lacity et al (1996) have argued that companies ought to be selective in deciding which systems or services to outsource and which to retain in-house. Following Loh and Venkatraman (1992b), Grover et al (1996) point out the need to take into account the degree of outsourcing as a continuous measure.

These propositions form the starting point of the theoretical development in this thesis. The main issues of concern is not so much whether to outsource or not but how to structure the organisation of IT supply. In this context, two dimensions have been established in the IS literature as being of prime significance. First, the degree of outsourcing as a continuous measure, determined by an array of criteria and trade-offs and second, the characteristics of individual systems and services and of

their interdependencies at a microanalytic level. Each of these dimensions has been considered in the literature in various useful but often limited and fragmented ways.

This thesis aims to bring together the parameters co-determining these two dimensions, under an integrative conceptual framework that could provide some more comprehensive guidance to the structuring of the governance of IT provision. Each of these two dimensions on its own gives a partial (and therefore incomplete) picture of the outsourcing decision. Focusing on selective outsourcing emphasises the distinct characteristics of systems and services but reduces the governance options to two unitary alternatives, namely outsourcing and insourcing. Focusing on the continuous array of governance possibilities addresses the structural details of contract and organisation but undermines the possibility of separate sourcing approaches for different services. Therefore, the two dimensions must be considered jointly. These basic ideas are fully articulated and developed in chapter 5. Chapter 2 sets the theoretical scene on which they are based.

Chapter Two

On the Economics of the Organisation of Production and Exchange

1. Introduction

By reviewing the relevant literature in the information systems field in the previous chapter, the main directions followed by this thesis in pursuing further research on IT outsourcing have been identified. The aim of this chapter is to review a range of theories in economics that can support theoretical development in these directions.

The emergence and proliferation of broad and long term contracts for IT systems and services caught IS academics and managers almost by surprise. However, such contracts have been very common, virtually since the dawn of the industrial revolution and much more so during the course of the twentieth century. Natural monopolies, as in the energy and transport industries, provide typical examples of long term, often exclusive, contracts. In many sectors of the economy, variousso called non-standard contracting practices, such as reciprocal trading, exclusive territories and resale price maintenance have been common.

Despite the abundance of such transactions and their obvious importance for economic activity, economists have paid relatively little attention to them. Traditional microeconomic theory treats firms as unitary transformers of inputs into outputs, analytically represented by the production function, without questioning how this is achieved. Vertical integration is often assessed for its informational properties (e.g. Riordan 1990). In a contractual relationship, supplier and customer have private information on the cost of production and the value of the product respectively. Under such conditions, if specialised investment

is required, profitable trade may not be realised as each party tries to manipulate his private information to maximise his information rent. The buyer will have to make do with standard products offered in the competitive market and thus forego the benefits of specialisation. Vertical integration has an efficiency justification inasmuch as it eliminates information asymmetry.

Following Adam Smith's legacy, competitive, decentralised markets have been considered as the main form of exchange. Any deviations have been either classified as undesirable monopolistic attempts or attributed to externalities and market failures. The boundaries of firms have been considered to depend on technological inseparabilities and scale and scope economies, derived from the production function. Within this theoretical mindset, Stigler (1968) extended Smith's conjecture that the specialisation of labour is limited by the extent of the market, to firm level and showed that as markets grow firms should become smaller and more specialised.

IT outsourcing as a phenomenon, can be attributed to these conditions. As the IT industry has been growing over time, extensive outsourcing should have been predictable. Moreover, it is possible that specialised suppliers may be able to achieve economies of specialisation, especially with the proliferation of standard techniques and technologies. Despite the plausibility of these suggestions, it can be convincingly argued that they are not adequate to explain the particular contractual and organisational arrangements employed. The purpose of this chapter is to review the relevant literature in economics that has grown substantially over the past thirty years and which, despite its incomplete state at present, can help generate substantial insight to the problem at hand.

In a landmark article, Coase (1937) questioned the necessity of technological conditions for the emergence of firms, arguing that technologically related assets can still be owned separately. He suggested that the presence of transaction costs in the economic system is the real cause. He postulated an imaginary world free of

transaction costs where the allocation of asset ownership and many other legal and commercial arrangements taken for granted today would not matter. However, in the real world, transaction costs abound.

They are the costs of running the economic system, much like friction in mechanical systems. Transaction costs represent the real resources economic agents dissipate to maintain viable and efficient relations of economic exchange. Transaction costs include, among others, the costs of writing, monitoring and enforcing contracts, of negotiating and of making provisions for uncertainty. In a broad sense, transaction costs are found both in markets as well as within firms. The coordination, supervision and other activities of management can be thought of as activities aimed at negotiating, monitoring and enforcing a different kind of contracts within organisations. Ouchi (1980) provides an interesting complementary conception of transaction costs as arising from the need for equity and reciprocity in economic exchange: "*A transaction cost is any activity which is engaged in to satisfy each party to an exchange that the value given and received is in accord with his or her expectations*" (p. 130).

Transaction costs do not add any productive value. They are a 'necessary evil' that rational economic agents are expected to try to minimise. In answering the question why firms exist, Coase suggested that for certain activities, the transaction costs of firm organisation are lower than the transaction costs of market exchange. On the other hand, Demsetz (1997) argues that the emergence of firms is fully explained in the traditional theory on account of the benefits of specialisation. Hierarchical structures of direction and control are necessary in order to deploy specialised knowledge efficiently (also see section 6, paragraphs 6-5 below). Even so, transaction cost arguments are needed to explain the structure and boundaries of firms, as well as the various governance structures under which market transactions take place. Transaction cost thinking sheds new light on various economic phenomena, including vertical integration and long term contracting.

Coase's (1937) ideas had an incubation period of nearly half a century before spawning a wealth of literature on the study of transaction costs as determinants of economic organisation. The study of team organisation by Cyert and March (1963) was pioneering in this respect. Blois (1972) studied 'close relationships' between firms involving exclusive or nearly exclusive trade and coined the term vertical quasi-integration. Goldberg (1976) was among the first to criticise the prevailing microeconomic and contracting approaches to economic exchange which were based on unlimited rationality and zero transaction costs.

One of the first analytical breakthroughs in dimensionalising transaction costs was made by Klein et al (1978). They define the notion of appropriable specialised quasi rents as the source of postcontractual opportunistic behaviour. The quasi rent value of an asset is "*the excess of its value over its salvage value, that is, its value in its next best use to another user. The potentially appropriable specialised portion of the quasi rent is that portion, if any, in excess of its value to the second highest valuing user*" (p. 298, emphasis in original).

When specialised assets are committed and such appropriable quasi rents created, a dependence arises where one party has an incentive (equal to the appropriable quasi rent) to take advantage of the other. For example, an IT supplier who trains a number of employees on a nearly unique information system of a particular client will have to sacrifice part of this investment in training if, for some reason, this contract is terminated too early and these employees have to be redeployed to another client where these skills will be only partly valuable. Once the investment is made, the client is in a position to seek further pecuniary or non-pecuniary gains commensurate to that portion of the training investment being at risk.

It is this that creates the problem of hold-up. The size of the hold up potential depends on three conditions: the presence (and size) of specific capital (i.e. appropriable quasi-rents), contract specification costs (i.e. the incentive for

contract violation) and contract enforcement costs (i.e. the ease of contract violation) (Klein 1980).

There are two kinds of governance responses to this problem, namely, vertical integration and long term contracts. First, the cost of devising sufficient explicit contracts is commensurate to the size of the appropriable quasi-rents. In addition, explicit contracts are plagued by rigidities, despite formal mechanisms to reinstate flexibility, such as price adjustment mechanisms (Crocker and Masten 1991). The rigidity of long term contracts defines a self-enforcing range, within which contingencies can be dealt with by the contract in an efficient or nearly efficient manner. A sufficiently large change in the environment (e.g. in demand, in the technology or in the market structure) forces the contract out of its self-enforcing range. The parties will have to renegotiate an appropriate response to the changed conditions. This creates further possibilities for hold up (Klein 1993). Second, the costs of vertical integration, however, are independent of the size of appropriable quasi rents. Therefore it is expected that over a certain (though varying) threshold level of specific investment and appropriable quasi rents, vertical integration will be the more efficient alternative. The promise of long term exchange operates as a future bond or insurance premium which has to be greater than the short term gains from opportunistic defection. Such a condition may not always be achievable by long term contracting. Internal organisation does not suffer from the same problem.

In general, this literature has (broadly) advanced the argument that, as asset specificity increases, economic activity is organised under more intricate governance structures (e.g. long term contracts) of which vertical integration is the 'final frontier'. However, recently Coase (1990) took exception with this proposition, arguing that increased levels of asset specificity do not necessarily explain vertical integration. By referring to the oft-quoted case of the acquisition of Fisher Body by General Motors in 1926 as evidence of the explanatory power of asset specificity, he juxtaposed the case of another GM supplier, AO Smith which

was the main supplier of frames to General Motors for many decades through a long term contract. Despite the highly specific investments required in tooling, production processes and human capital, the two companies never integrated and managed to trade efficiently over a long period of time and across a long distance.

Two main questions arise from this critique. First, why does the cost of organising particular activities differ among firms. Second, if internal organisation is so efficient in mitigating opportunistic hazards, why isn't the whole economy better organised as a single large firm. Economic theory has not given a definite answer to the first question. A simple response is to attribute such differences to situational contingencies. It can tentatively be argued though that given the same technologies and skills, the differential costs of organising economic activity can be ascribed to all the detailed elements comprising the notion of governance. Economists have produced numerous models of different aspects of the organisation of exchange. Each of these models can offer a partial answer to the question.

The second question can be restated as 'what is responsible for the limitations to firm size?' or, in comparative terms, 'why can't a large firm do everything that a collection of small firms can do and more?'. The response is twofold. First, there are a number of economic activities for which transaction costs are relatively limited so that market organisation can serve them more efficiently in comparison to the hierarchy. Second, the size of firms is limited by the perils of bureaucratisation (Putterman 1995). Despite this paramount significance of bureaucratic failures, it has received relatively little attention. Influence costs (Milgrom and Roberts 1992), incentive slack and the impossibility of selective intervention (see section 3.3 below) are postulated as sources of bureaucratic costs.

This brief sample of transaction cost thinking demonstrates the intricate complexities arising from the analysis of transaction costs. The purpose of this chapter is to clarify the main tenets of the literature with the aim of setting an

economising basis for understanding IT outsourcing arrangements, as opposed to political, sociological or other perspectives.

Williamson (1985) places the intellectual locus of transaction cost economics on the efficiency branch of the study of economic organisation, in sharp contrast to the monopoly arguments that pervade the traditional microeconomic literature (Figure 2.1). Moreover, transaction cost economics is distinguished from the incentive branch of the efficiency approach in that the latter emphasises precontractual alignment of divergent goals using incentives, while the former focuses on postcontractual structures and processes for adaptation. The incentive branch can be further split into the property rights literature and the agency theory literature. Transaction cost economics also forks into the measurement and governance strands. Despite the differences between these four strands of the literature, they have a lot in common and, on the whole, they usefully inform each other.

The structure of this chapter broadly follows this classification of the literature. First, though, a brief note on the legal and institutional context of economic exchange provides a useful point of reference. Then, transaction cost economics, agency theory and the theory of property rights are discussed in turn. Two important and recurring themes are then presented, namely, the role of knowledge and information and the effects of culture and reputation. The concluding section provides a critique of the literature and sets certain further assumptions and propositions used in subsequent chapters.

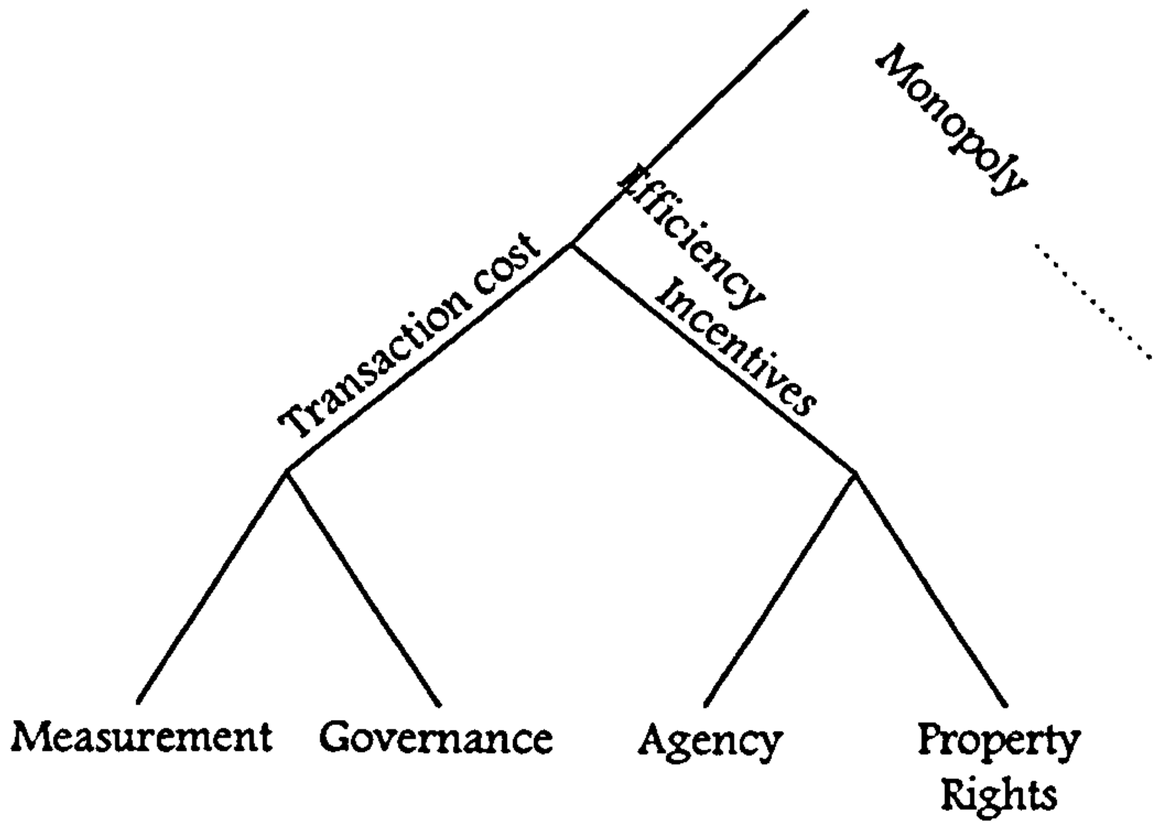


Figure 2.1

A Classification of the Literature on Economic Organisation

Amended from Williamson (1985: 24).

2. Institutional environment: the transaction and the firm

It is widely accepted that the law has important ramifications for economics and vice versa (Goldberg 1989). Much more so in the study of transactions, where notions such as contract, court enforcement, litigation and arbitration are, among others, the objects of study. In a general sense, the law is one part of the institutional environment and enables economic activity to take place by delimiting the possible actions of economic agents and enforcing equity where it is not privately achieved. The institutional environment entails numerous formal and informal organisations, ranging from the state and trade organisations to the enforcement of property rights (e.g. through patents) and the power of reputation. An explanation of the institutional environment (see North 1990) is out of the scope of this brief note. Instead, a clarification of the main legal positions on exchange and transactions and on the concept of the firm will prove useful in understanding the theories subsequently reviewed.

2.1. Relational exchange

Exchange is a direct result of specialisation. If individuals were entirely self sufficient there would be no need for exchange. The ubiquity of exchange can, therefore, be thought of as a result both of the efficiency and wealth benefits of specialisation as well as of the introduction of money. In addition, trade has desirable wealth effects in its own right. It has to be recognised though, that humans depend on each other not only for materialistic wealth maximising purposes, but also for broader social and psychological 'satisfaction'. Consequently, economic exchange has to be seen as embedded in a web of social relations, despite the fact that these are not directly studied by economic analysis.

Classical contract law and neoclassical microeconomics typically treat transactions as discrete exchanges, isolated from any historical, social or psychological context. Discrete exchange is defined as "*relatively free of relations beyond those created by a common language, a system of order [...], a monetary system, and, for discrete exchanges not accomplished simultaneously, a legal system enforcing promises*" (Macneil 1985:485, emphasis in original). Examples include purchasing petrol on the motorway or music CDs by mail order.

Contracts are interpreted as promises of the parties involved, to be fulfilled into the future. Such contracts are supposed to presentiate[1] all future possibilities and provide appropriate responses in advance. This contracting paradigm relies on the efficacy of courts to enforce promises or award remedies and damages according to some rules of justice.

However, contracts are necessarily incomplete. Presentiating all future possibilities (contingent claims contracting) is ultimately impossible. Essentially, there is not one point in time when the parties to a contract make a definite and final agreement.

---

[1] Presentiation is the term used to denote that contracts in a way map future events and actions onto the present by means of an agreement of promises.

Instead, it is an incremental process of renegotiation and adaptation over time. Actual contracts very often allow many decisions to be made during the course of the relationship rather than in advance. Litigation is costly and not always efficacious as the courts have limited knowledge of the context of the exchange and many claims the parties make cannot be directly verified, not to mention delays and possible miscarriages of justice. As a result, private ordering is often more efficient and the threat of litigation is employed to delineate threat positions in bilateral bargaining. Llewellyn (1931) succinctly summarises: "*a contract almost never accurately indicates real working relations, but... affords a rough indication around which such relations vary, an occasional guide in cases of doubt, and a norm of ultimate appeal when the relations cease in fact to work*" (p.737, quoted in Williamson 1996:10).

Moreover, the historical and social context as well as the identity of the parties play a significant role in determining the content and execution of contracts. Cultural and reputation effects within specific industries, ethnic communities or geographical regions testify to this. In addition, the parties to a contract (whether long term or repeated spot transactions) develop reciprocal social and psychological links beyond the letter of the agreement during the course of the exchange relationship. Macaulay (1989) quotes research sowing that managers tend to disregard contractual technicalities and prefer to rely on 'gentlemen's agreements'. Lowry's (1976) comment on the social conditioning of economic exchange is apposite: "*[...] the bargaining transaction [assuming it leads to actual trade] is essentially divisive and zero sum, and [...] it must occur in a contained zone structured by a relational system. On the other hand, the relational structure necessary to an organised economy cannot be derived from an aggregation of bargains, but must be understood as a social contract developed in a setting of widely accepted social values and goals.*" (p. 18)

Moreover, Macneil (1985) strongly argues that wealth maximisation is not the only goal that people pursue through their economic transactions, particularly

within relational transactions. On non-pecuniary satisfaction in economic exchange, Williamson (1975:38-39) notes that "*recognition that alternative modes of economic organisation give rise to differing exchange relations, and that the relations themselves are valued, requires that organisational effectiveness be viewed more broadly than the usual efficiency calculus would dictate ... efficiency and a sense of well-being (which includes, but transcends, equity) are intrinsically (nonseparably) joined ... individuals [may] forego material gains for nonpecuniary satisfactions*". Whitford (1985) suggests that identity (in the sense that humans, as social entities, construct their identities partly through the relations they form) and participation are two other goals pursued by individuals, independently of any wealth effects of their transactions.

This is the meaning of relational contracting. It is a broad (in many respects sociological) notion of exchange as primarily human relation in which individuals seek to satisfy a number of goals through an incremental process of mutual adaptation to changes of all kinds. The law makes appropriate provisions in support of relational contracting. Macneil (1978, 1985) argues that all contracts are ultimately relational to a greater or lesser extent. Even the most discrete exchanges are also embedded within a broader relation, such as the market that enables their performance. The extent to which an exchange is discrete or relational is a matter of degree along a continuum, rather than kind.

2.2. *The firm as a discrete alternative to the market*

Arguably, the distinction between the firm and the market is more a matter of degree than kind. It is a matter of the degree of discretion, authority, continuity of association, incentive intensity and so on. These properties clearly vary from spot markets, through intermediate governance structures, to long term contracts and internal organisation. The discrete differences often ascribed to internal organisation in terms of its transaction cost economising capacity in a general sense, can be disputed (Alchian and Demsetz 1972). Direction and authority exists

within firms as well as it exists in markets. Firms can direct the actions of their employees and can fire them, as they can direct the performance of their suppliers and terminate their supply relationship. Information asymmetries are also not relieved with integration, as employees can still withhold, distort or not produce information. Commercial contracts can also enforce audits, as firms can use fiat to obtain information. Similar arguments can be articulated with respect to incentives and adaptation. This has lead many to argue that the firm is no more than a legal fiction. It is a nexus of explicit and implicit contracts (Aoki et al 1990) and the difference between the firm and the market lies on the particulars of each contract.

The agency theoretic position of Jensen and Meckling (1976:311) in this respect has been a point of reference for many arguments: "*it makes little or no sense to try to distinguish those things that are 'inside' the firm and those that are 'outside' of it. There is in a very real sense only a multitude of complex relationships (i.e. contracts) between the legal fiction (the firm) and the owners of ... inputs and the consumers of output*". This has been the position adopted in many economic theories, for two apparent reasons. First, it is a convenient analytical concept. Second, any discrete differences between market transactions and the firm certainly do not derive endogenously from the assumptions and analysis within the relevant theories. Even when Williamson (1996) insists on the structural discreteness between markets and hierarchies, it is an assertion (or assumption) rather than a result of transaction cost theory.

Masten (1993) argues that this position is probably correct until one brings into consideration the legal regimes governing market transactions and the firm[2]. As it turns out, the law provides substantially different mechanisms to the firm and the market, thus rendering them discrete alternatives indeed. The law provides much broader power to the authority of employers than the performance that might be

---

[2] The work of Masten (1993) and Clark (1985) reviewed here makes reference to American legislation. Although labour laws may differ in other jurisdictions, the main principles reviewed in this section are expected to be similar, at least among the western capitalist societies.

stipulated in the employment contract. Clark (1985) locates the legal differences between the firm and the market in four legal attributes of the fiduciary relationship within the firm. First, the affirmative duties to disclose. Second, open-ended duties to act. Third, closed-in rights to positional advantages. Fourth, moral rhetoric. These are briefly explained below.

An employee is expected by labour laws to obey all reasonable rules and orders of the employer, to yield loyalty, respect and faithfulness and to disclose all relevant information. In return, the employer is held liable for any harm the employee causes to third parties in the course of their employment. This creates the need for closer monitoring but also an incentive for the employee to follow the employer's orders without much consideration of the consequences. In a commercial contract the contractor is only expected to abide by the performance stipulated in the contract and has complete control over all non-contracted aspects of his behaviour. Moreover, the law affords the employer greater powers to enforce the cooperative behaviour of the employee. The employee can be held liable and the employer can recover damages for failure to perform the duties accruing from the law and the employment contract, though not necessarily explicitly stipulated in it. In addition, members of the firm are explicitly forbidden to use any privileged information (positional advantage) to their private benefit (e.g. insider trading). Of course, this creates an incentive for the employee not to seek additional valuable information thus increasing the importance of monitoring. Finally, termination is the ultimate threat for both commercial and employment contracts. The differences mainly lie in what constitutes breach of contract in each case. An employer can take various disciplinary actions or dismiss an employee on the basis of disloyalty, dishonesty and insubordination, among other 'offences' which are irrelevant in the context of commercial contracting.

While labour laws may differ between jurisdictions, in general the employment relation is granted a greater capacity of authority, flexibility and information advantages as compared to commercial contracts. Of course, many such aspects of

internal organisation can be replicated by contractual agreement. However, writing a contract that would fully replicate internal organisation would be extremely costly and therefore impractical. Moreover, the more complex a commercial contract the less flexible it is by definition. The legal essence of the firm, therefore, constitutes a governance structure combining the above mentioned advantages at a very low cost while achieving substantial adaptive flexibility. In addition, the legal constitution of the firm also shows where the incentive deficiencies commonly ascribed to the firm lie. The firm is therefore a discrete structural alternative to the market by reference to the legal regimes that govern each institution. For many analytical purposes, however, it is still possible to conceive a continuum from the spot market to the hierarchy, while recognising that the switch from long term contracting to internal organisation makes a substantial difference in kind and not only in degree, despite the many structural similarities between the two.

3. Transaction cost economics

The origins of transaction cost theory can be traced back to Coase (1937) while substantial theoretical advances have been made by Williamson (e.g. 1975, 1985, 1996). Transaction cost economics emphasises the institutions of governance which refer to alternative forms of organising economic activities. These are contrasted to the institutional environment which refers to the broader legal, political and customary 'rules of the game' in a historical context (e.g. North 1990). The institutional environment is exogenous to the study of the institutions of governance and is considered as a set of shift parameters affecting the feasibility, efficacy and relative costs (efficiency) of alternative governance mechanisms.

Transaction cost economics has been applied to a number of issues, including non-standard contracting practices that are normally assessed for their anticompetitive effects, the employment relation, corporate finance and governance and public policy. However, it is typically considered in connection with the question of 'make or buy'.

The main argument is that transactions, which differ in their attributes, are (or should be) aligned to governance structures, which differ in their incentive and adaptive capacities, in a discriminating and transaction cost economising way. Moreover, it is recognised that all governance structures are ultimately imperfect. Consequently, the selection and assignment of governance structures is made on a comparative rather than an absolute basis (Williamson 1996).

The following brief exposition of transaction cost economics starts with the basic behavioural assumptions of the theory and their implications. Then the dimensions of transactions and governance structures are outlined. Finally, the process of alignment is sketched.

3.1. Behavioural assumptions

The theory of transaction cost economics is based on two fundamental behavioural assumptions about human nature, namely, bounded rationality and opportunism.

Economic actors are "*intendedly rational but only limitedly so*" (Simon 1961). Both the intentional character of human rationality and its limited nature are crucial assumptions in this theory. All the theories of the firm reviewed in this chapter rely on at least a minimum amount of rationality to justify the efficiency perspective. The necessary incompleteness of contracting (which creates all the contractual hazards that transaction cost economics is interested in) is a direct result of limitations to rationality (combined, of course, with environmental uncertainty and measurement problems). Once it is recognised that all feasible contracts are necessarily incomplete, those mechanisms of governance that can deal with *ex post* manifestations of this incompleteness (such as the condition of small numbers, bilateral dependence or hold up) are of primary importance. These mechanisms (including contingent claims, renegotiations or litigation) bear various costs which have to be judged against their efficacy. That rationality is bounded does not mean

that it is myopic. On the contrary, individuals have to be farsighted to acknowledge contract incompleteness and to anticipate future hazards, despite the fact that they cannot determine them in advance.

Economic agents are also characterised by opportunism, defined as self interest seeking with guile. Opportunism gives rise to various well known agency problems such as adverse selection and moral hazard (see section 4). Both these problems appear under conditions of hidden and/or asymmetric information. If economic agents were not acting opportunistically they would normally share all available information to maximise total net gains. Because of opportunism, they prefer to appropriate any surplus not accounted for by extant agreements and thus strategically manipulate (withhold, distort or selectively disclose) their private information. This behavioural attitude is so strong that agents often withhold their private information not realising that sharing it might increase their individual gains (prisoner's dilemma). The fundamental implication of opportunism and its manifestations for economic organisation relates to the typical notion of contract as promise. Opportunistic agents cannot be assumed to reliably fulfil their promise. Various costly *ex ante* and *ex post* measures (such as incentives, monitoring or court ordering) are needed to counter this possibility.

Williamson (1975) argues that information impactedness is a derivative condition of bounded rationality and opportunism. Pre- and postcontractual information impactedness are distinguished and the role of first mover advantages with respect to this condition is recognised. Information impactedness arises either from asymmetrically available information between two parties, coupled with high costs of achieving information parity and the opportunistic behaviour of the parties, or from identically informed parties who, however, make opportunistic representations that the true state of affairs is something different to what they both know, coupled with the high costs of apprising an arbiter of the true information. Information impactedness is a particularly acute problem under 'small number' conditions where the absence of a competitive process allows opportunistic

behaviour (otherwise, competitive bidding and benchmarking serve both as revelation mechanisms and as deterrents of opportunism, despite the added problems of under-bidding). When information is incomplete *ex ante* and the parties cannot devise sufficient rules of conduct, *ex post* haggling results as each party dissipates resources to take advantage of opportunities to increase their private gains.

In this respect, the purpose of economic organisation, according to transaction cost economics, is to economise on bounded rationality and to safeguard against opportunism. It is also worth mentioning that transaction cost economics, as advanced by Williamson, typically ignores the risk attitudes of contracting parties and implicitly assumes risk neutrality. In order to effect an alignment of transactions and governance structures, given these assumptions, the dimensions of transactions and governance structures need to be operationalised. This is dealt with in the following two sections.

3.2. Dimensionalising transactions

Transactions can be described along three main dimensions, namely, frequency, uncertainty and asset specificity.

Recurrent or long term transactions, as opposed to one-off exchanges provide fertile ground for most of the phenomena of economic organisation. Incomplete contracting, opportunism and reputation effects are some of the main conditions that become more relevant in recurrent or long term transactions. One-off exchanges approximate more closely the notion of discreteness while recurrent or long term transactions provide more ground for the development of complex relations (see section 2.1).

Uncertainty is of three types, namely, primary, secondary and behavioural (Williamson 1985: 56-60). Primary uncertainty is of a state-contingent type and

derives from natural or social events out of the control of the decision maker. Secondary uncertainty is a result of bounded rationality and refers to distorted communication, limited access to information and measurement problems which are a result of limitations to the cognitive skills of individuals. When such problems are a result of strategic manoeuvring (e.g. non-disclosure, disguise or distortion of information), they are due to opportunism and the resultant uncertainty is called behavioural.

Asset specificity refers to the extent to which an asset cannot be redeployed to alternative uses, except at a significant cost or sacrifice of productive value. It is related to the notion of sunk costs but whereas in neoclassical economic analysis sunk costs have no role to play after they are committed, asset specificity takes a central position in transaction cost analysis. The degree of asset specificity is commensurate to the appropriable quasi-rents it creates. There are four main kinds of asset specificity, namely, physical asset specificity (usually involving specialised equipment or materials), human asset specificity (including tacit knowledge and learning by doing), site specificity (spatial proximity normally due to technological constraints) and dedicated investments (large investments in general purpose assets made for the benefit of a particular trading partner in anticipation of high volume future exchange). In addition, brand name capital and temporal specificity have recently been added as further types of asset specificity (Williamson 1996: 59-60). Zaheer and Venkatraman (1995) introduce the notion of business process asset specificity as deriving from the combination of human asset specificity and procedural specificity, the latter referring to the degree to which a firm's workflows and processes are customised in line with the requirements of the exchange partner.

A profound implication of asset specificity is what Williamson calls the 'Fundamental Transformation'. While competitive bidding is typically efficacious at the outset, once a contract is entered into by the initial winning supplier and the customer, competitive parity at contract renewal is questionable. If the exchange

relationship does not involve any specific assets of any kind, all potential suppliers, including the incumbent, remain at a par (assuming that the cost of bidding and supplier selection is negligible). Otherwise, the incumbent contractor has an advantage over other potential suppliers, commensurate to the unredeployable value of the specific assets involved (appropriable quasi-rents). This is the 'small numbers' condition that arises postcontractually.

By definition, the presence of asset specificity implies that, were the contract to be terminated, economic value would have to be sacrificed. It is therefore in the interest of both parties to renew their contract (or to avoid termination due to conflict). For rational decision makers therefore, it is a question of balancing the costs and potential benefits of switching. This line of thought is particularly pertinent to the outsourcing of IT services where a variety of practices is being observed, from companies trying to sustain their relations with incumbent suppliers despite the conflicts, to other organisations painfully switching suppliers, or others going as far as re-establishing their internal capability all over again.

3.3. *Dimensionalising governance structures*

The main purpose of economic organisation according to this theory is to align transactions and governance structures in a discriminating, transaction cost economising way. To achieve this, we need an operational means of characterising transactions and governance structures. While frequency, uncertainty and asset specificity are the dimensions characterising transactions, three main attribute categories dimensionalise governance alternatives (Williamson 1991).

Williamson identifies three discrete governance structures, namely markets, hierarchies and hybrids, although most comparisons are typically made between the first two extremes. The three attribute categories of these governance structures are (i) instruments, including incentive intensity and administrative controls, (ii)

adaptive capability, including autonomous and cooperative adaptation, and (iii) contract law regimes, including classical, neoclassical and forbearance doctrines.

Incentive intensity is highest in markets and lowest in hierarchies, despite the fact that the latter also have access to additional incentive instruments, such as performance payments, profit sharing and career rewards. Administrative controls are the fundamental instrument of hierarchies, exemplified by authority and fiat. The threat of termination and possible reputation effects are the only such instruments in market transactions. The power of authority in hierarchies is constrained by its limited capacity to monitor and process all the necessary information. In markets, whatever contingency is not adequately dealt with by the price mechanism, is resolved in the courts of law. Alternatively, the parties involved switch to another governance structure (hybrid or hierarchy) where more administrative flexibility is available (examples of such recourse include take-overs and more generally vertical and lateral integration).

According to Williamson, adaptation is the central problem of economic organisation. There are two kinds of adaptation. First, autonomous, where markets have an advantage and second, cooperative, where hierarchies have the advantage. Intermediate (hybrid) contractual forms replicate some degree of cooperative adaptation within what is essentially a market transaction.

Markets effect adaptation via the autonomous decisions and actions of economic agents reacting to the information conveyed by prices. Any signs of cooperation are either condemned as anticompetitive (collusion) or justified on grounds of market failure, externalities and/or transaction costs. In contrast, hierarchies adapt to changing circumstances by cooperation, or the "*conscious, deliberate and purposeful efforts to craft adaptive internal coordinating mechanisms*" (Williamson 1996: 103). In pure hierarchies, autonomous adaptation is not, in principle, available as an option, although divisionalisation, transfer pricing or other attempts to decentralise control are efforts to replicate some autonomous

adaptation of the type available in markets. That adaptations in hierarchies are called cooperative does not necessarily imply the presence of congenial relations and congruous motives. Opportunism survives within hierarchies as it does in markets, despite the provisions made by the law (see below and section 2.2). As a result, the adaptations actually implemented are not necessarily efficient, as members of the organisation pursue their individual subgoals. Inasmuch as authority in hierarchies does not only intervene to supplement incentive and adaptive inadequacies, truly selective intervention is impossible.

Markets are governed under classical contract law regimes. Classical contract law attempts to facilitate exchange by emphasising discreteness: "*sharp in by clear agreement, sharp out by clear performance*" (Macneil 1974:738) and presentiation (fully pre-specifying all relevant future contingencies and responses pertaining to a given transaction). The exchange agreements (explicit or implicit) and the legislation describe all possible actions available to economic actors. Any disputes are resolved by the courts according to the relevant agreements and laws. Contingent claims contracting is the respective concept in economics. In a departure from his earlier analyses, Williamson has replaced relational contracting (1985) for the forbearance doctrine (1991) as the legal regime governing the operation of internal organisation. In sharp contrast to classical contract law, the forbearance doctrine allows maximum flexibility to the firm to regulate all its internal affairs by its own mechanisms, without even the possibility of bringing a disagreement to the courts. Exceptions to forbearance refer to issues creating externalities, usually relating to employee rights, health and safety and the natural environment.

Hybrid governance structures possess all elements of the three attribute categories available to markets and hierarchies, but to a more limited extent. The parties to a hybrid exchange relationship maintain their ownership autonomy, thus high-powered market incentives may still be effected, albeit to a restricted degree. Hybrid contracts afford substantial administrative controls (for information disclosure,

joint decision making, renegotiation, dispute resolution etc.), in recognition of the problems arising due to contract incompleteness, reciprocal dependence and asset specificity. Inasmuch as parties to a hybrid governance structure maintain their autonomy while instituting joint administrative controls, they sustain both autonomous as well as cooperative adaptation. Neoclassical contracting is the legal regime governing hybrid contracts. Faced with the impossibility of complete *ex ante* contingent claims contracting, neoclassical contract law reinforces exchange agreements with administrative mechanisms allowing *ex post* resolution of contingencies.

3.4. *Aligning transactions with governance structures*

Williamson (1996: 92) argues that the principle according to which transactions are assigned governance structures in a discriminating way is that of remediableness, whereby first-best outcomes are infeasible due to the ultimate incompleteness of all feasible contracts. Consequently, the least costly alternative is selected on a comparative basis.

Moreover, Williamson (1996) emphasises discrete structural analysis of governance structures (as opposed to continuous variations in governance attributes) on three grounds. First, it is argued, the three generic forms (market, hybrid and hierarchy) employ different means and are not extensions of each other. Second, each of them relies on a different contract law regime. Third, Williamson focuses on first order economising as opposed to marginal analysis which is concerned with second order refinements (e.g. in incentive intensity or information acquisition).

Williamson (1985) and Riordan and Williamson (1985) have attempted a formalisation of the bipolar choice between market and hierarchy, taking into account the degree of asset specificity and firm output. The production and transaction cost differentials between internal organisation and market

procurement are studied under various amounts of asset specificity and firm output. For relatively small amounts of output and limited asset specificity, market procurement has an advantage in terms of production costs (arising from scale and scope economies achieved by specialised suppliers). This advantage diminishes as output grows and the firm can replicate similar production economies. In terms of transaction costs, the market has an advantage for low levels of asset specificity and the hierarchy has an advantage for higher levels of asset specificity.

For a given amount of output, there will be a threshold level of specificity for which the hierarchy replaces the market as the least cost alternative. Williamson (1996) argues that contracts in the region of that threshold are expected to be ambiguous and unstable: "*mixed governance, in which some firms will be observed to buy, others to make and all express 'dissatisfaction' with their procurement solution are apt to arise*" (p. 69). One cannot avoid pondering whether this is what we currently observe with IT outsourcing.

The presence of economies of scale and scope favour market organisation over a wider range of asset specificity values than would be observed, if production cost economies were absent. Having said that, Teece (1980) argues that the mere presence of scope economies (Panzar and Willig 1981) is not a necessary nor a sufficient condition for lateral integration, unless the relevant assets are also co-specialised.

Another ambiguity arises when, within a range of limited output, the marginal production cost disadvantage of internal organisation and the marginal transaction cost disadvantage of market procurement operate in opposite directions. At a sufficiently high level of output, internal organisation can replicate all production cost economies available in the market, thus the governance choice rests on the governance cost differential between the two alternatives, for a given level of asset specificity.

This treatment, although instructive and consistent with the analytical standards of the neoclassical tradition, is limited in two main respects. First, it deals with two polar extremes, vertical integration and market contracting, without considering the particular structure of any of the two which determines the extent of governance costs and of the respective differentials. This is a result of the insistence on first order economising and on discrete alternatives. Second, it examines only one transaction at a time, whereas contracting takes place in a broader institutional nexus (Aoki et al 1990) of possibly related exchange relationships.

Another similar exercise in developing the analytics of the alignment of governance structures with transactions is presented by Williamson (1991). This includes three governance forms (market, hybrid and hierarchy) and ignores production costs. The transaction cost curves of the three generic governance alternatives are expressed as an increasing function of asset specificity. Due to the adaptability differences discussed earlier, the three curves differ in their slopes and origins. The market turns out to be the least cost option for low asset specificity but its cost increases rapidly with specificity. Hybrid forms are the least cost alternative over a medium range of asset specificity. The governance cost curve for the hierarchy has the highest origin and the smallest slope and turns out to be the least cost structure for high levels of asset specificity. An efficient (or governance possibility) frontier of governance structures is thus created, whereupon actual governance structures (e.g. franchising) can be placed.

The threshold values of asset specificity for which a transition from markets to hybrids and from hybrids to hierarchies takes place are affected by the institutional environment. Four elements of the institutional environment are examined by Williamson (1991) in this respect, namely, property rights, contract law, reputation effects and uncertainty. Weak enforcement of property rights (e.g. an ineffective patent system) would shift all governance costs upwards, particularly as asset specificity increases. Other things being equal, hierarchy would be expected to be the least cost option for a broader range of asset specificity. Changes in the

willingness of courts to intervene (i.e. changes in the contract law regimes) would also shift the governance cost curves of the respective governance structures, thus altering their relative efficiency. An environment where reputation effects are effectively communicated and play an important role, reduces the costs of hybrid governance and render it feasible over a broader range of asset specificity. Increased uncertainty reduces the efficacy of hybrid forms of governance while hierarchies can rely on fiat and markets on autonomous adaptation to respond to contingencies.

To further demonstrate the role of asset specificity as the main dimension of transactions and the main determinant of governance, Williamson has proposed a simple contracting schema with three main contracting possibilities. First, transactions that are efficiently supported by general purpose assets do not need protective governance structures. Discrete market contracting suffices and competition obtains. Second, transactions that involve significant transaction-specific investments typically involve some sort of bilateral dependence. Such transactions can be supported by contractual safeguards to protect against postcontractual opportunism and to provide incentives for the (presumably more productive) specific investment. Third, transactions involving specific assets but for which no contractual safeguards are provided will lead to a higher price. For example, to induce a supplier to make a transaction-specific investment, the buyer will have to pay for that part of the investment that the supplier cannot salvage outside the particular relation. Moreover, such a transaction is expected to be contractually unstable as the pricing-out of specificity may not mitigate other contractual hazards arising from opportunism, uncertainty and measurement problems. Such transactions are expected to revert to either a position where the specific investment is forgone in favour of the general purpose assets or to a position with appropriate governance safeguards. Such safeguards provide credible commitments to the exchange relationship, so that the appropriation of the specific investment is guaranteed and opportunistic dangers avoided.

The provision of credible commitments is analogous to an exchange of hostages (Williamson 1986). One possibility when cost-reducing specific investments are needed, is for the buyer to make the investment and retain ownership, while contracting with the supplier for the supply of goods using these specific assets. The feasibility of this option, however, is limited to physical and mobile assets, and creates further incentive distortions (dampened performance and possible abuse of the assets).

Another alternative is for the buyer to promise (post a hostage of appropriate value) to compensate the supplier for the unrecoverable cost of the specific investment, should the exchange come to an end, while otherwise pay the marginal cost price. This option is still subject to supplier opportunism. The supplier may induce breach of contract in order to appropriate the hostage. For this reason it is suggested that buyers post hostages which are of the appropriate value to them (thus ensuring credibility of commitment) but are of no value to the supplier (thus mitigating the risk of contrived cancellation).

Further problems arise when the buyer cannot verify the extent and value of specific investments claimed to have been made by the supplier. For this reasons contracts often provide for the engagement of third party arbiters and/or give the buyer audit rights over the supplier's activity. Such audit rights appear to be typical in IT outsourcing contracts.

Another source of governance hazard arises due to the necessary incompleteness of contracts and the opportunity of the parties to haggle over contingencies that arise, especially where the allocative implications are not straightforward. Specialised contractual arrangements allow hybrid forms of governance to control such hazards. Third party arbitration and reciprocal exposure (perhaps through reciprocal trading) are two of the possibilities. Other possibilities include joint decision making mechanisms and procedures of private appeal (escalation

processes). The promise of long term exchange on the part of the buyer is also a form of hostage.

This discussion is particularly pertinent to IT outsourcing which typically involves a substantial amount of specificity of almost all kinds. One example is proprietary software. This constitutes physical asset specificity, which also extends to human asset specificity, since specialised skills are needed to develop and maintain it. Another example is desktop support which requires investment in local operations (site specificity). Vendor staff working exclusively on a particular client's project or service, constitute a dedicated investment, to the extent that the vendor cannot employ them efficiently, were the particular contract to be terminated. The case of Gloucester County Council (chapter 1, section 1) is illustrative of the hold-up potential accruing to both dedicated investments and human asset specificity associated with data centre operations. Long-term agreements are one commitment mechanism typically employed to safeguard specific investments.

4. Agency costs

The literature on agency costs is divided into two parts (Jensen 1983). First, mathematical principal-agent theory (e.g. Pratt and Zeckhauser 1985) emphasises *ex ante* alignment of incentives, usually assuming unlimited rationality, costless and complete contracting and perfect commitment. Positive agency theory (e.g. Jensen and Meckling 1976) studies both *ex ante* alignment and *ex post* adaptation and adopts the main tenets of transaction cost theory, namely that contracting is costly and incomplete. Therefore, organisational form is of prime importance for this branch of the theory while for mathematical principal agent theory, organisation form does not matter as virtually any outcome is achievable through an efficient (no matter how complex) contract.

Both theories view the firm as a nexus of internal and external contracts or as a network of overlapping or nested principal-agent relationships between owners, managers, employees, suppliers and customers. Mathematical principal agent

theory structures these contracts through mechanism design whereas positive agency theory defines the nature of the firm as an equilibrium state of this nexus of contracts.

The contract is the object of study and the individual the unit of analysis in both strands of the theory. The presence of specific investments (appropriable quasi-rents) as a cause of *ex post* hold-up (opportunism) potentials is also a central theme. However, mathematical principal agent theory can almost always perfectly resolve hold-up potentials contractually (Rogerson 1992) by relying on the absence of externalities (one individual's decisions and actions do not affect others), risk neutrality, costless contracting (contracts can be arbitrarily complex) and perfect commitment (individuals cannot renege or renegotiate, even if this is *ex post* efficient).

Information asymmetry is recognised as another source of opportunistic behaviour which manifests itself in two ways. Precontractual information asymmetry creates the problem of adverse selection. For example, a buyer has limited information on a potential supplier's capacity to deliver high quality and/or low cost. The supplier can consequently take advantage of the buyer's ignorance. Postcontractually, information asymmetry leads to moral hazard. For example, the contractor may shirk or exaggerate his costs or any adverse conditions, if the buyer cannot monitor these variables sufficiently. Moreover, if such information is not independently verifiable, the buyer may renege on the contract and refuse to make payments by understating the contractor's performance (moral hazard from the part of the principal).

Positive agency theory is positioned much closer to transaction cost analysis (Williamson 1996: 171-179) because they share the same behavioural assumptions (although agency theory admits bounded rationality only implicitly by recognising contract incompleteness) and they both appreciate the importance of

transaction costs in the broadest sense of the term. The following sections outline the main tenets of each branch of the theory.

4.1. Mathematical principal-agent theory

The basic setting (Sappington 1991) involves a principal contracting with an agent for a particular job. The principal makes an offer which includes a payment schedule conditional on the performance of the agent. The payment is equal to the principal's valuation of the agent's performance minus a fixed franchise fee. The agent accepts the contract if his expected gains exceed his opportunity cost. The principal motivates the agent by making him the residual claimant of any performance that he might achieve above the fixed franchise fee. The agent's performance can be deterministic or probabilistic, depending on his effort and a random parameter representing the environmental conditions that may arise.

The typical principal-agent setting assumes complete and symmetric information, complete contracts, risk neutral agent, perfect commitment from both principal and agent and costless measurement and verification of information. While transaction costs are typically ignored in this branch of the literature, differential risk attitudes often play a significant role.

If the agent is risk averse and the principal risk neutral, the latter will choose to bear some of the risk himself, thus offering a compensation based more on realised costs rather than the agent's performance. The agent thus ceases to be the sole residual claimant, which dampens his incentives. Moreover, this insurance provided by the principal can be a cause of moral hazard. The same result obtains when the principal has to respect the agent's limited liability (limited commitment) and has to compensate his effort even under adverse environmental conditions. The certainty with which effort leads to performance in a particular task environment determines risk. By opting for input (effort) monitoring (and input-based

compensation), the principal internalises these risks and vice versa (Anderson and Oliver 1987).

More generally, if the principal is the residual claimant of gain and risk, he will prefer to monitor inputs. Characteristic examples include vertical integration and cost plus contracting. If the agent is the residual claimant, the principal will prefer to measure output performance (e.g. fixed price or performance contingent contracts). The argument can be reversed so that if input monitoring is more efficacious (or less costly) the principal will prefer to acquire the residual claims and if output measurement is more efficacious (or less costly) the principal will choose to assign residual claims to the agent. These results derive from the dual sources of rent available to agents, namely, manipulating input costs and quantity (Khalil and Lawarree 1995). Effort (input) monitoring can generally be used in addition to performance (output) measurement, even if either information is imperfect, as long as the total information is more useful. The feasibility of all these options is ultimately constrained by the risk attitudes of each party, which are typically set as assumptions in principal-agent models.

Team organisation poses additional problems for input monitoring and output measurement, depending on the degree of separability (the extent to which output information discriminates between different input owners) and programmability (the availability of knowledge on the transformation of inputs to output). To relieve this problem, two or more competing agents can be employed as reciprocal monitors, as long as each agent's effort and performance is independent of the other's. Agent collusion then becomes a problem which can, in turn, be mitigated if each agent is rewarded for truthfully squealing on the other. Collusion between principals and agents can also occur in multi-tiered or nested agency relationships, such as those found in hierarchies, for example between managers, supervisors and workers (Tirole 1986).

Depending on risk preferences, simple tournaments can provide powerful incentives. The principal does not have to measure cardinal performance but offers 'winner-take-all' or 'loser-bear-all' contracts where only ordinal performance needs to be assessed. When a small amount of effort is only required, 'loser-bear-all' contracts that penalise non-performance are more efficacious. Such tournaments are also effective deterrents of opportunism (reneging) on the part of the principal, assuming that ascertaining a winner is less costly and more verifiable than assessing the cardinal performance of each agent.

Under symmetric information, greater environmental uncertainty is expected to lead to a greater fixed and a smaller incentive proportion in compensation, and vice versa. When the principal has private information about environmental uncertainty and the effort level required by the agent does not vary with (uncertain) changes in the environment, the principal will choose to offer a contract with a higher incentive proportion as a signalling device to communicate the degree of uncertainty to the agent (Umanath et al 1996). If the agent's effort level is expected to vary with states of the environment, then the incentive effect is expected to dominate (over the signalling effect) and compensation will follow the rule under symmetric information. If the agent has private information, the principal can induce truth revelation by offering a menu of alternative reward schemes, more or less incentive intensive. The agent will have to select the payment structure that maximises his gains given his private information and thus he will not be able to expropriate any information rents by misrepresenting his private information. If risk preferences and the availability of information about them are taken into account, matters get more complicated and the above mechanisms may, under risk aversion, be rendered ineffective altogether.

Contracts for software development are typically of the cost-plus kind (time and materials), in recognition of the uncertainties involved in the process. Large outsourcing contracts often have a mixed payment schedule where a (relatively smaller) proportion of total compensation is tied to performance targets.

Interdependent tasks have further incentive implications. When the costs of two tasks are negatively correlated, better incentives are provided if they are both assigned to the same agent and vice versa. For example, if higher development cost (effort) leads to lower maintenance cost, a countervailing consideration (Riordan and Sappington 1987) is introduced to the agent who is assigned both. Reporting high development costs implies a claim that maintenance will be carried out at low cost. Alternatively, an agent undertaking only development has an incentive to exaggerate his cost and so does the agent separately undertaking maintenance. However, if, for example, high costs of requirements specification leads to high cost of development (positive cost correlation), there is no countervailing incentive and the agent can exaggerate both costs (or, alternatively, shirk at both tasks). The principal is therefore better off assigning the two tasks to separate agents.

In long term contracts, if future gains are not heavily discounted, the principal can offer adequate incentives while protecting the agent from uncontrollable risk by making payment contingent on average performance over time meeting some specified target. This assumes that the technologies and the environment are fairly stable over the averaging time window. Averaging offers some protection to the agent against unpredictable contingencies impairing his performance.

The threat of termination in long term contracts has powerful incentive consequences (Sen 1996). Termination clauses are instituted for three main reasons of efficiency. First, if the incumbent contractor is somehow assessed as inferior to alternatives, termination provisions allow the principal to switch. Second, in the presence of moral hazard (i.e. when the agent's effort is not perfectly monitored) the threat of immediate termination upon detection of defection can induce agents to align their actions with the interests of the principal. Finally, termination clauses can be used to deter the agent from misrepresenting his private information (e.g. costs). Such a termination clause should specify that, other things being equal, the agent should be fired if he reports a high level of costs. The efficacy of this

mechanism depends on the extent of intertemporal cost correlation, or, alternatively, the degree of uncertainty. Negative or no cost correlation (high uncertainty) reduces the efficacy of this kind of termination clause.

In other words, this applies when, for example, the agent performs a repeated task and has private information about his costs. If the task environment is stable, the agent should be fired when he reports unprecedented high cost. This rule is weakened when the task environment is known to be uncertain and/or when the agent's private information can be monitored or verified.

4.2. Positive agency theory

Jensen and Meckling (1976) define the agency relationship as a contract under which one or more persons (the principals) engage another person (the agent) to perform some service on their behalf which involves delegating some decision making authority to the agent. Both parties to the relationship are assumed to be self-interested utility maximisers. This means that the agent will often make decisions and take actions that diverge from the interests of the principal. The reduction of welfare experienced by the principal due to this divergence is the source of agency costs. To limit this cost, the principal will offer incentives aiming at realigning the interests of the agent with those of the principal. Moreover, the principal will incur monitoring costs aiming at limiting the incongruent activities of the agent and informing the amount of compensation (including possible penalties) that should be awarded. The agent will also often expend resources (bonding costs) to offer guarantees that he will be acting in the interest of the principal. However, both monitoring and bonding efforts are imperfect so that the remaining divergence will still create a loss of welfare. The sum of the principal's monitoring costs, the agent's bonding costs and the residual loss, defines agency costs.

Positive agency theory has been extensively applied to corporate governance but has broader implications. A recurrent theme is that of the separation of ownership from control (Fama and Jensen 1983). Ownership is defined as residual risk bearing or alternatively, residual claim on net gains. Moreover, four stages of the decision making process are distinguished, namely, initiation, ratification, implementation and monitoring of decisions. Initiation and implementation combined refer to decision management, while the combination of ratification and monitoring refers to decision control.

The benefits of specialisation in decision making and risk bearing and the need to control the agency costs arising from the separation of ownership from decision making determine the concentration or separation of residual risk bearing, decision management and decision control. The benefits to specialisation in decision making are defined by the task environment and by the amount of specific (non-transferable) information available to individual agents. Accordingly, Fama and Jensen (1983) formulate two fundamental hypotheses.

First, separation of residual risk bearing from decision making leads to decision systems that separate decision management from decision control. This is necessary in order to control agency costs. Second, combination of decision management and decision control in a few agents leads to residual claims that are largely restricted to these agents. This leads to efficiency losses in risk bearing. The second arrangement is likely to be found in cases when specific information is limited to these few agents. In complex organisations where specific knowledge is diffused to many individuals, residual risk bearing is expected to be separated from decision making. To counter the agency problems created as a result, decision management is separated from decision control. Incentive mechanisms provide important support in linking the interests of owners, decision managers and decision controllers. In certain circumstances, usually when there are few residual claimants, decision control will be assigned to them. However, when residual claimants are many and diffuse (and cannot exercise effective control collectively), separation and diffusion

of decision management and control helps mitigate the power of individual agents to expropriate the interests of residual claimants. Formal decision hierarchies is the organisational structure that often obtains. This discussion applies directly to the relations between shareholders (risk bearers), the board of directors (decision controllers) and executive directors (decision managers).

Based on Fama and Jensen's account alone, the common practice in IT outsourcing contracts where the contractor (who is responsible for decision management) is also asked to monitor and report performance (decision control), appears as an anomaly. To the extent that the contract specifies a cost-plus reward scheme, the client is also the residual claimant (the contractor is the residual claimant to an extent equal to the proportion of performance contingent rewards in his compensation). This apparent anomaly can be justified on three factors. First, the client organisation (principal) is also the recipient of the service and as such can costlessly collect performance information and exercise control. The quality of the information and the cost of its collection are compromised by the fact the this information is diffused across the users of the organisation and is not directly observable by the individuals managing the contract. Moreover, some performance information is more rarely observed (e.g. security breaches). Nonetheless, even this imperfect signal can, for many services, provide a sufficient deterrent to the supplier's interest in misrepresenting performance information, as long as appropriate penalties are attached to unacceptable performance. Second, anecdotal evidence suggests that reputation effects can be quite powerful, particularly for large, high profile contracts. Third, there may be substantial benefits from specialisation to monitoring achieved by the contractor (reduction of measurement cost). The rights to audit and/or benchmark, often reserved by clients, also serve in part as additional monitoring instruments. Additionally, the client organisation often retains a portion of decision management by reserving the right to approve certain decisions made by the supplier. There is no doubt, however, that all these means are imperfect, hence some degree of opportunistic expropriation does remain (residual loss) as this theory predicts.

A further incentive problem arises when performance targets are regularly reset according to past performance. *Ex post* efficiency dictates that principals should not commit to previously agreed performance targets if newer information from the execution of the contract can be used to revise incentives. Such lack of commitment gives rise to the ratchet effect (Litwack 1993). An agent anticipating this behaviour will underperform in order to avoid the imposition of more stringent incentives in the future. To avoid this adverse incentive effect, the principal should offer assurances that as long as the technology, the task environment and the agent(s) are the same, incentive targets will not change, even if the originally agreed targets prove to be too high or too low. Alternatively, introducing some randomness in the task environment (e.g. by job rotation or new technology) can mitigate the ratchet effect while efficient adjustments can take place.

According to the notion of the firm as a nexus of contracts, agency problems are also pervasive within hierarchies. Holmstrom (1989) succinctly demonstrates how the promotion system which is the traditional incentive mechanism of internal organisation may misallocate specialised human resources. The structural nature of hierarchy creates further distortions arising from the strategic manipulation of private information (McAfee and McMillan 1995). This cost increases cumulatively as the distance in hierarchical levels from the employees who 'own' local information to those who use it for decision making widens. This result justifies in part flatter hierarchies, allocation of decision authority to line employees (empowerment) and subcontracting.

Influence costs are the resources dissipated in organisations by the efforts of individuals to take advantage of their private information in order to influence those with decision making authority to their own benefit (Milgrom and Roberts 1990, 1992). Two conditions are necessary to render influence costs likely. First, there must be a group of decisions that determine how the costs and benefits are allocated within an organisation. Second, the parties affected must have open

channels of communication to the decision makers during the time period when decisions are being made. The trade-off that arises is between minimising influence costs while maximising the communication of useful information for decision making. In other words, the goal is to structure the decision making process in such a way as to limit opportunities for rent seeking while still acquiring the necessary information to make a well informed decision. The total size of influence costs is commensurate to the amount of rents available for redistribution. The propensity of an individual or organisational unit to engage in influence activities depends on their ability to expropriate information rents. Decentralisation and disaggregation of decision making and information generation is one means of limiting influence activities by allocating decision authority to the owners of relevant information.

5. Incomplete contracting and property rights

In the absence of transaction costs, economic activity can be organised efficiently both in markets and in hierarchies without the allocation of ownership making any difference. Any rights that naturally accrue to asset ownership can be costlessly reassigned as dictated by efficiency, through contracting. However, transaction costs abound in reality, making complete contracting impossible. Uncertainty, measurement problems and bounded rationality are responsible for contract incompleteness.

That contracts are incomplete means that many actions over the assets involved in a transaction cannot be contracted in advance. These residual rights of control that are not allocated by contract or by law define ownership (Grossman and Hart 1986, Hart 1993). Contract incompleteness also implies that some part of the profit stream generated by the use of an asset cannot be contracted in advance. Ownership determines the *ex post* bargaining power of the parties to a transaction over the division of these rents. Moreover, the presence of transaction specific investments creates potential *ex post* hold ups. Inasmuch as ownership, defined as residual rights of control, affects *ex post* bargaining and hold ups, efficient

allocation of ownership should minimise *ex post* bargaining costs and help mitigate *ex post* hold up potentials. As a result, the allocation of ownership affects the parties' *ex ante* incentives to invest in specific assets.

Residual rights of control are sharply distinguished from residual claims over future profit streams which feature prominently in agency theory. In general, residual rights and residual claims do not have to be allocated to the same party. For example, an employee or contractor who works on a profit sharing incentive scheme is to an extent the residual claimant but not necessarily the owner of the relevant assets. There are at least three reasons, however, why the two should go together. First, the separation of residual rights and residual claims intensifies problems of opportunism. Second, residual claims tend to strengthen incentives for short term profit, possibly at the expense of the asset's long term value, while the incentives accruing to ownership redress this balance. Third, separation may not be feasible altogether if contracting for the allocation of profit streams is imperfect. The owner of residual rights of control ultimately controls residual claims as well (Hart 1995).

Hart and Moore (1990) advance the implications of this theory for vertical integration. Some of the main results are summarised here, assuming that there are two economic agents, a buyer A and a supplier B, initially each owning asset *a* and asset *b* respectively. If agent A's investment in a transaction is more important than agent B's, agent A should acquire B, otherwise they should remain independent. Agents A and B should remain independent when agent B is indispensable to asset *b*, when asset *b* is idiosyncratic to agent B and/or when assets *a* and *b* are economically independent. Agents A and B should integrate when agent A is indispensable to asset *b*, when asset *b* is idiosyncratic to agent A, when assets *a* and *b* are complementary and/or when agent B is dispensable (in the sense that he can be readily replaced without any loss of value). Clearly these results are stylised and rather limited, as multiple assets, multiple agents and a combination of the above conditions are the more general case, also likely to be more common in practice.

The theory extends to such cases, specifying various control structures including, for example, shared ownership.

According to this theory, when an investment specific to the buyer needs to be made then ownership should remain with the buyer. Specifically, Hart (1993:152) argues that an asset which is indispensable to a buyer, who, in turn, can transfer it to contracts with alternative suppliers while it is of less value to the supplier outside the particular contract, should remain under the ownership of the buyer. Supplier ownership of this asset would lock-in the buyer and give the supplier the opportunity to hold him up. In this sense we would expect that clients maintain ownership of IT assets in IT outsourcing deals. The observed practice of transferring all assets to the supplier is either a clear contradiction to this theory or requires more elaborate examination of the facts that might justify it.

There are three ways to reconcile this apparent anomaly. First, if the performance required from the supplier by the contract involves taking actions that cannot be explicitly specified in the contract (e.g. innovation, process improvement or replacement of equipment), then performing such actions requires possession of the residual rights to control the assets, or otherwise every such action would have to be negotiated separately. Such negotiations can become overwhelmingly costly. Therefore, inasmuch as all the activity of the supplier cannot be contracted *ex ante*, transfer of ownership (as residual rights of control) provides incentives for such activities to take place. Second, asset ownership creates incentives for the supplier to make further client-specific investments of his own. Third, IT outsourcing contracts typically transfer asset ownership to the supplier but specify their return to the buyer, after appropriate payments are made, at contract termination. The only apparent obstacle to virtually costless contracting of asset ownership restitution is the calculation of their remaining value. The supplier is in a position to bargain over the remaining value of assets and can also manipulate his own accounting methods. In response, standard accounting techniques and third party auditing can be specified in the contract.

This theory works for physical assets of various kinds, including 'hard' assets such as machinery and production plants, as well as 'soft' assets such as patents, client lists, existing contracts or brand name. In the absence of slavery, residual rights of control over human assets are held by each individual and are not transferable. Nonetheless, the importance of human assets is paramount. The theory assumes that individuals acquire or invest in skills and specialised knowledge specific to a given physical asset. This is how the conditions of indispensability and idiosyncrasy mentioned above accrue.

However, the theory also insists that the firm cannot exist without any physical assets at all. It makes an even stronger claim, that "*control over non-human assets leads to control over human assets*" (Hart 1995: 58). That a given physical asset is idiosyncratic to an agent means that the agent needs this asset in order to be productive. If this agent is a contractor and the exchange relation breaks down, the agent takes the asset with him. However, if this agent is an employee, termination of the employment relation would deprive the agent from the asset he needs in order to be productive. In this case he is likely to be more loyal to the employer. This is how the property rights theory of the firm explains the greater power of authority and fiat often ascribed to internal organisation.

The incomplete contracting theory of the firm works at the level of the individual owner-manager. However, the diffusion of ownership among many shareholders and the percolation of control rights down a hierarchy are still amenable to analysis. Moreover, it is assumed that individuals have adequate finances to purchase the assets as efficiency dictates. In reality, of course, various investors, bankers or shareholders may be involved and thus the efficient allocation of ownership and control rights becomes a more complicated matter. Several developments have been made in this respect but are out of the scope of our interest in this chapter.

The discussion on residual claims and on diffused ownership (among shareholders and investors) and control (among managers) indicates that ownership is not a binary variable but it can take intermediate values (Hart 1995). This is an underdeveloped area of the theory, possibly due to the problem of studying residual control rights directly, since, by definition, they cannot be specified.

6. Measurement costs and the firm as knowledge repository

As Barzel (1982) points out, measurement costs are pervasive in economic activity. They arise from private information and its opportunistic manipulation, from hidden physical attributes of commodities, from team organisation and from specialised non-transferrable expertise, to name but a few. Their presence explains the emergence of organisational arrangements as diverse as warranties, product bundling and block booking, professional associations or share contracts.

Based on the measurement problems of team organisation, Alchian and Demsetz (1972) advance the informational properties of internal organisation to one of the more influential discussions on the nature of the firm. Team organisation is defined as the joint production by two or more inputs where the joint output is greater than the sum of the outputs if each input was used in separate production. The main problem identified is that of measuring (and subsequently controlling) the marginal productivities of each input. Due to the systemic nature of team organisation, measurement is costly and imperfect and thus each team member has an incentive to shirk and free-ride on the rest of the team.

Market competition cannot by itself alleviate this problem. A competitive market of team member 'replacements' would be operative only if productivity and shirking information was costlessly available. Instead, firm organisation institutes a monitor whose main task is to observe the behaviour of team members, rather than their output, and use this information to ascertain their productivity. Making the monitor the residual claimant of the net gains of total team production provides

adequate incentive for him to perform efficiently. Overall, the monitor is also given responsibility to be the central party contracting with each member (more generally, for each input), to alter team membership and to sell all his rights and responsibilities. This, according to Alchian and Demsetz (1972), defines firm ownership (or the employer) of the classical capitalist firm.

The informational and systemic view of the firm is further analysed by Aoki (1990) who argues that an important feature of internal organisation is information (knowledge) that is generated, maintained and used collectively by organisational members and is not individually portable by them (as opposed to specialised skills). Such information creates information rents, which, due to their participatory nature, are non-contractible. Such information rents only accrue to the collective activity within the particular organisational framework and both the collective nature and the organisational framework are indispensable to their generation. Inasmuch as such conditions cannot be replicated by contractual agreement, when participatory information rents are important parameters of performance (or elements of final net gains), internal organisation is the preferred governance structure, *ceteris paribus.*

This argument is extended by Demsetz (1993) who follows Adam Smith's original ideas and the refinements offered by Stigler (1968), that knowledge accrues to specialisation and specialisation can, in turn, take place either within or across firms. The organisational arrangements within firms, particularly the process of directing, serve to economise on the generation, accumulation and use of specialised knowledge.

Specialised knowledge can be put to work without loss of its value, first by directing the activity of non-experts to produce and second, by selling products (or services) which encapsulate this knowledge and require less information to use than to produce. Hierarchical structures of direction economise on the costs of acquiring and maintaining specialised knowledge, as fewer people need to specialise in an

area of expertise (or, alternatively, a greater amount of knowledge can be handled by the same number of people). The transfer of an intermediate product relieves the downstream purchaser from the need to acquire and use the same knowledge as the upstream producer. The need to economise on the specialised knowledge required for production and to achieve the production economies accruing to this specialised knowledge, determines firm size. The limits of the firm are reached when the breadth of knowledge required to direct a range of operations exceeds the capacity of the firm, assuming there are knowledge economies to scale up to a certain volume of knowledge and diseconomies thereafter. In other words, the division of knowledge between the direction of employees and the direction of buyers of the good determines the boundaries of the firm.

The transaction cost implications of organisational knowledge, particularly under conditions of change, uncertainty and/or innovation (Wieland 1993), create second order effects for the determination of firm boundaries. For example, the downstream purchaser of an intermediate product may be able to create innovative knowledge regarding the design or construction of the intermediate good. However, he may have inadequate incentives to do so if non-contractibility that accrues to the collective nature of the innovation prevents him from securing property rights (Lundgren 1990). Backward integration may thus obtain.

Although the informational properties of team organisation and the creation of participatory information rents have been examined in the analytical economic tradition, the notion of the firm as knowledge repository has not received as much analytical treatment. Winter (1993) reaffirms the systemic nature of the firm by postulating that knowledge and the associated human asset specificity are to be found in the collectiveness of firm members and in organisational structures and processes which are more than the sum of individuals. However, there remains a need to unpack this notion into an account of the exact structures and processes by which this knowledge is generated and preserved while individual members of the firm come and go.

This part of the literature may find fertile grounds for empirical work in the organisation of IT systems and services. It can be hypothesised (see also chapter 1) that the growth of the number of different systems, standards and technologies makes it difficult for a firm to acquire and maintain all the relevant expertise, particularly when faced with frequent innovations in the IT industry.

7. Culture, reputation and trust

The firm-specific common stock of knowledge is often defined as corporate culture (Cremer 1990). Besides any anthropological and psychological facets of culture, for economists it has efficiency advantages as it helps economise on bounded rationality and to a certain extent it conditions the incentives of organisational members (Zaheer and Venkatraman 1995).

Kreps (1990) defines corporate culture as a principle or a set of principles that guide the decisions and actions of the members of the organisation when unforeseen contingencies occur. As such, culture is primarily a remedy for contract incompleteness. It instils an idea as to how the organisation as a whole and each member individually should react to circumstances as they arise. Therefore, culture is also a coordination device, inasmuch as complete centralisation of decision making and comprehensive communication of directions are ultimately infeasible. A set of shared tacit principles (culture) relieve the need for comprehensive communication and coordination, thus facilitating the exercise of hierarchical authority.

Culture is also a control and screening device. Members of the organisation can be monitored for their adherence to the principle(s). Non-members can be screened as to their 'compatibility' with the organisation's culture, should they become candidates for membership or other long term association.

Although culture evolves over time to accommodate the kinds of contingencies the organisation tends to encounter, it is (or it has to be) relatively more persistent than individual organisational members and individual decisions or plans. For culture to be operative it has to be common knowledge in a recursive sense. Each member of the organisation is aware of it and is also aware that all other members are aware both of the culture and of the truth of this statement. Therefore, it has to be communicated and 'taught', especially to new members, through the continuous and consistent application of its principles. This means that the principles of culture will have to be applied even when, *ex post*, they are not the most efficient reaction to a given contingency. This is necessary in order to reinforce and communicate it, both to insiders and to outsiders. Moreover, the particular circumstances that rationalise a deviation from the principles of culture may not be common knowledge to all relevant parties, so that the deviation may be interpreted as the demise of culture and thus dampen their incentives to attend to it. In turn, this observation points to the fact that culture is largely indivisible, in the sense that the principles that comprise it are strongly interrelated.

Finally, Kreps argues that culture has implications for firm size and scope. The argument is similar to that for organisational knowledge in the previous section. Culture is specialised in the sense that it has evolved to encounter the kinds of contingencies typically faced by the given organisation. As the size and scope of the firm increase, so does the range of different types of contingencies that culture will have to deal with. This can be achieved in two ways. Either more specialised principles are added or the same core principles are applied to the broader set of contingencies. In the first case more ambiguity is introduced as to which principle should apply and in the second, responses are more inappropriate, more often. Therefore, there is a limit to the size and scope of the enterprise that a coherent and consistent culture can sustain.

The implication of the above discussion for inter-firm transactions, is that the parties' cultures will have to be common knowledge (in the strict, recursive sense).

In addition, there will have to be some kind of compatibility in the sense that contingencies should be dealt with in non-conflicting ways. Otherwise it would be difficult to reach agreement as each party would rightly insist in maintaining their own principles, as discussed above. Culture defines a significant part of a firm's identity.

It is often argued that corporate culture is an extension of the reputation mechanisms that pervade economic activity (Williamson 1996: 156-157). Reputation effects also rely on the condition that the rewards and penalties of honouring or taking advantage of trust respectively, are common knowledge. However, the principle underlying reputation is a generic notion of trust (see further below) rather than a specialised set of historically determined principles. Still though, reputation effects serve to counter opportunism in the absence of complete contracting.

In a one-off exchange, one party may decide to trust another and expose itself in anticipation of jointly beneficial trade. The second party is often better off betraying the other's trust and appropriating any benefits, if future exchange is not anticipated or not valued (discounted). In the context of repeated transactions, however, both parties are better off honouring each other's trust in anticipation of further future trade and knowing that abusive behaviour is penalised at least by loss of repeat trade. This argument can be extended to a broad setting with many buyers and sellers where each party essentially inherits the reputation of their predecessors.

Such reputation effects are pervasive but are also limited in a number of respects. A given outcome (payoff) from a transaction may be probabilistic and if one party cannot directly observe the other's behaviour it will be uncertain as to whether a bad outcome is due to the other party's defection or to unfortunate circumstances. The efficacy of reputation mechanisms is therefore limited by noisy and indirect observations.

Incomplete contracting implies that in some contingencies it will not be clear what constitutes abusive or cooperative behaviour since appropriate responses are indeterminate *ex ante* and haggling is likely to result. Reputation information is also often imperfectly communicated and, moreover, it is imperfectly understood.

Further distortions are introduced if goods and services and, consequently, anticipated performance, change frequently over time or when agents in a competitive environment manipulate reputation signals strategically. The effectiveness of reputation mechanisms relies on the enforcement of penalties to those parties that defect. In many cases, penalties for defection are not enforced and the effectiveness of the mechanism is degraded. Often the costs of termination (if this is the penalty) are higher than those of a temporary defection. As long as the relationship is not terminated, both parties have an incentive to transmit positive reputation signals to the market in order to maintain their relationship, even in the presence of serious conflicts. In addition, a given party may plead for mercy claiming that it should not be held responsible for the bad reputation of its predecessors and achieve forgiveness. Alternatively, a certain party may ignore information of bad reputation, if it considers itself to be 'smarter' and thus more able to achieve the cooperation of others. Complexity may introduce further problems.

Despite these detrimental side-effects, trust and reputation are necessary to economic exchange. Trust accrues to personal relationships (not necessarily of an economic kind), to the institutional environment (e.g. the state, the legal system, the local or ethnic community) and to the transactions of economic agents. To the economist, trust is based on calculativeness while humanitarian or religious conceptions of trust cannot be accommodated. Even in personal trust, where altruism admittedly prevails, there has to be a (low) threshold of calculativeness in the sense that a rational individual is not expected to endure too many betrayals of his trust (Williamson 1993).

In this context, trust can be defined as a subset of risk, as the subjective probability with which an agent assesses that another agent will perform a task that is beneficial or at least non-detrimental to him. Reputation informs the calculation of that probability (trust). According to this conception of trust, governance safeguards (such as hostages - see section 3.4) are properly interpreted as manifestations of trust (since they increase the probability of a net gain), despite the more intuitive approach according to which in the presence of trust such safeguards are redundant.

Calculativeness is therefore the pervasive basis for such trust. It is the result of opportunism combined with bounded rationality. Admittedly though, calculativeness can be carried to extremes and has to be suppressed. Effectively, this is a calculative limitation of calculativeness, a second order calculativeness, in itself (Williamson 1993).

The literature on IT outsourcing, particularly early commercial publications, emphasised the importance of concepts such as trust and partnership in total outsourcing agreements. This hyperbole over trust has, in many instances, lead to an interpretation of trust as substitute to a good contract. The discussion in this section helps to put matters into perspective. Trust and convivial corporate cultures between client and vendor are valuable on account of the necessary incompleteness of such contracts. Both qualities, however, are (or should be) calculated and must lead to appropriate contractual safeguards aimed at limiting the potential scope of future opportunism (from either party).

8. Conclusions

Hierarchical organisational forms are ubiquitous in modern, market economies. Simon (1991) uses the hypothetical analogy of a Martian who, upon approaching earth, colours the organisation of economic activity on a map of the developed

economies. The Martian would find the largest areas of the map covered by hierarchical organisation with fewer areas covered by market transactions. The emergence of firms, however, is not foreign to the neoclassical economic theory (Demsetz 1997). Firms emerge as a result of economies of specialisation. Demsetz further argues that a reduction in transaction costs does not lead to the replacement of hierarchies by markets, instead, it increases both the number of firms (of a smaller average size) and the number of market transactions in the economy. However, it is widely agreed that neoclassical economics has relatively little to say about the boundaries and structure of firms and contracts. The theories briefly reviewed in this chapter address these issues head on.

Among these theories, alternative postulates have been put forth, regarding the motivation of firms in expanding their activities vertically and laterally and the factors that might limit the ability of firms to grow in these directions. Further, studies of some of the internal phenomena of hierarchical organisation have also been examined. All these theories provide efficiency explanations of observed structures, while monopoly explanations have been set aside.

Three further omissions of this review ought to be admitted at this point, namely the evolutionary theory of economic change (Nelson and Winter 1982), the resource-based view of the firm (Pfeffer and Salancik 1978) and the X-inefficiency theory of the firm (Leibenstein 1975). The first introduces in a systematic manner the importance of history in economic organisation, as opposed to a comparatively static solution to the problem of organising. It has been bypassed precisely because the aim of this research is, in part, to suggest a normative framework that management can employ in answering the question of how to organise the provision of IT services. However, by no means does this aim lessen the significance of the evolutionary antecedents of the IT outsourcing phenomenon as it emerged in the late 1980s, a study of which remains pertinent.

The resource-based view of the firm is related to the evolutionary theory and emphasises the competencies and capabilities that organisations develop through learning. Moreover, the resource dependence between an organisation and its social and economic environment is an important theme in this theory. The discussion on knowledge and culture in the previous two sections is more closely related to this perspective. Otherwise, economic theories such as transaction cost economics are often presented as competing (Pfeffer 1997: 192-193) perspectives. However, it is conceivable that transaction cost economics, agency, resource based and evolutionary theories can usefully complement each other. In particular, both empirical research and abstract theories on governance structures often incorporate constructs and arguments from more than one of the above perspectives (e.g. Anderson and Oliver 1987, Yarbrough and Yarbrough 1988, Alston and Gillespie 1989, Donaldson 1990, Lieberman 1991, Mahoney 1992, Zaheer and Venkatraman 1995, Roberts and Greenwood 1997, Hodgson 1998 and Nordhaug 1998). Moreover, Williamson envisages the convergence of all these theories into a unified theory of organisation, or what he calls 'the incipient science of organisation' (1990). The fact that it is not clear exactly how all these theories might be unified, should not inhibit us from taking advantage of the benefits of combining insights from them. This thesis addresses the structure of IT outsourcing contracts by drawing on all the theories reviewed in the present chapter.

The theory of X-inefficiency takes an even more microanalytic perspective as compared to, for example, transaction cost theory. It redresses several standard assumptions (e.g. utility maximisation) of more mainstream theories and provides an analytical framework for the study of the behaviour of economic agents as individuals with rich personality traits. However, this theory has remained at the margin of intellectual development in the economics of organisation.

Tentative evidence (particularly in conjunction with that provided in chapter 1) shows that the theories reviewed in this chapter are pertinent to the study of the organisation of IT services. Clearly, all these theories are far from complete. Each of

them, taken individually, illuminates many important aspects of the phenomena under study. Furthermore, they appear to offer complementary explanations in several respects, despite the marked differences that distinguish them.

A growing literature of empirical work, has found supporting evidence, particularly for transaction cost and agency theory (e.g. Lafontaine 1992, Joskow 1993, Lyons 1994, Zaheer and Venkatraman 1995, Masten 1996). In addition, the apparent complementarity between these two approaches has prompted several conceptual and/or empirical studies that combine them (e.g. Ouchi 1980, Yarbrough and Yarbrough 1988, Mahoney 1992, Zenger and Hesterly 1997). It is becoming clear that such a combined approach, despite the complexity and other difficulties it entails, is a profitable research direction. Such empirical work, however, faces significant difficulties. Most of the concepts advanced by the above theories, such as asset specificity, uncertainty and residual rights are not always directly observable or measurable in the same way as prices and quantities. Faced with such difficulties, empirical work is based either on case studies or on econometric analyses using various imperfect approximations (see for example Rindfleisch and Heide 1997).

The approach followed in this thesis departs from the mainstream methodology of economics. Some fundamental theoretical assumptions are discussed in the following section. Methodological issues and goals are the subject of the next chapter.

8.1. The theoretical basis of this thesis

The analysis in subsequent chapters is based on the fundamental arguments of transaction cost economics, as put forward by Williamson, supplemented by insights from agency theory, property rights, knowledge and reputation effects. These diverse elements are brought together by attempting a critique and

reinterpretation of the basic transaction cost arguments which, in turn, makes the necessary space for these other theoretical contributions.

Transaction cost theory has been presented in the literature in a fairly comprehensive manner, including fundamental assumptions and arguments as well as operationalisations and applications, into a tightly knit system of ideas. For this reason it provides a convenient basis on which to lay the principles of a combined theory of organisation. The other theories examined have been presented in numerous, relatively more narrow accounts, although they often go to greater analytical depths and, of course, they do not lack broad applicability. The greater analytical depth is precisely the feature that can complement the basic principles of transaction cost thinking.

Transaction cost theory claims to be microanalytic in that it focuses on the transaction as the main unit of analysis. In turn, transactions take place "*when a good or service is transferred across a technologically separable interface*" (Williamson 1996: 58). The study of transactions should therefore focus on individual products or services. Having said that, the questions of how aggregate an individual good or service is and where technologically separable interfaces are placed, are open to interpretation. An obvious example is that 'IT service' is usually not a unitary service but includes a variety of diverse sub-systems and sub-services. The composition of this bundle is crucial as different sub-services may be characterised by varying degrees of asset specificity, measurement costs and inter-relationships with other systems and services. Given a sufficiently broad population of systems and services that a company may employ, the specification of such bundles is non-trivial.

Transactions are also often identified with contracts, if there are any. However, the correspondence between goods or services and contracts is not necessarily one to one. A single contract may include several goods or services. Similarly the correspondence between contracts and suppliers is not one to one. A buyer may

purchase a range of goods from a single supplier, under one or more contracts. This is one reason why a relevant microanalytic strategy ought to focus on a finer level of detail, examining the composition of the goods or services being exchanged. By aggregating diverse components and assigning a generic value of 'mixed asset specificity' the discriminating power of the theory is compromised and other important dimensions (such as incentives, measurement issues and technical interdependence) are overlooked.

Further, Williamson contends that transaction cost economics is concerned with discrete structural alternatives, namely markets, hybrids and hierarchies, and goes to great lengths showing that they represent discrete differences in kind rather than continuous variations in the degree of certain parameters. This argument is backed with an emphasis on "*first order economising (getting the basic alignments right) rather than second order refinements (adjusting the margins)*" (Williamson 1996). A discussion of the latter is deferred until certain reasons why discrete structural alternatives are not sufficient, are first explained.

It is difficult to generate a normative theory of organisation, which is also going to be useful to the practitioner, with just the three alternatives proposed by Williamson. It is not adequately clear what action a manager should take if he is advised that a certain transaction must be organised under a hybrid governance structure.

Williamson (1996: 90) quotes a conversation he had with Ian Macneil in the late 1970s, speculating on the frequency distribution of markets, hybrids and hierarchies. Macneil suggested a bell-shaped distribution with most governance structures being of the hybrid kind, whereas Williamson posited a bimodal distribution, with most transactions split between markets and hierarchies, arguing that transactions in the middle range are inherently unstable.

In another instance, however, Williamson (1985: 385) argues otherwise: "*... it is sometimes argued that models of polar extremes are wholly adequate. But the relevant test, presumably, is whether middle-range phenomena can better be understood and refutable implications derived by studying these matters directly. As matters stand at present, contracts in the middle range are notoriously intractable. But if that is where the main contracting action resides, more attention to mixed transactions is arguably warranted*".

Although the frequency distribution of governance structures remains an interesting question for empirical investigation, it can be argued that aggregating all 'middle range phenomena' under the 'hybrid' label is inadequate.

At a conceptual level of analysis, it is argued that pure markets and hierarchies are ideal types that never occur in practice. Even the purest of (actual) spot markets, such as foreign exchange markets, are closely controlled and regulated. To give another example, occasional purchases from a greengrocer's are also conditioned by reputation effects and the problems of measuring quality. More generally, once incomplete contracting is accepted as a universal condition of the actual world, pure markets are ultimately impossible.

Pure hierarchies are also an ideal extreme, expected to be rare in practice, on three accounts. First, hierarchies can encapsulate internal markets for goods and services (e.g. profit centres and transfer prices). These structures are akin to usual markets, despite the fact that they eschew the enforcement powers of courts. Second, and in relation to internal markets, is the capacity of hierarchies to accommodate autonomous adaptation (Williamson 1996: 110). The same transaction cost arguments explaining the multidivisional form (M-form) and other alternative structures of internal organisation suggest that hierarchies are not pure from elements typical of market structures (Williamson 1985: 279). Finally, inasmuch as exit (resignation) is an option, the reach of fiat is limited.

Zenger and Hesterly (1997) have convincingly demonstrated the dilution of market structures with hierarchical elements and the diffusion of market characteristics within hierarchies. Once it is thus established that pure markets and hierarchies are fictional ideal types, it follows that the frequency distribution of governance structures cannot be bimodal in the sense described by Williamson above. Having said that, establishing that pure markets and hierarchies are fictional ideal types, does not rule them out as good instruments for exposition, discourse and comparison. Moreover, having found elements of hierarchy within markets and elements of market within hierarchies does not mean that they are similar structures, with a small distance between them. On the contrary, hybrid governance structures are a very broad category occupying the space between markets and hierarchies.

Coase's (1993) argument that a high degree of asset specificity is not a sufficient condition for vertical integration, can be reversed and interpreted that vertical integration is not the only response to rising transaction costs. There is a multiplicity of governance mechanisms available to economic agents and a variety of circumstances encountered in practice, suggesting the existence of many variations of actual governance structures. Contracting practices in the middle range are contingent upon the actual circumstances which render certain governance responses more efficacious than others. Reputation effects, hostages of various kinds, ownership allocation, explicit incentives and integration are some of the governance alternatives which may prevail, individually or in combination, under different circumstances. In a pragmatic sense, each actual governance structure is unique. Much theoretical power is foregone by grouping together an array of different governance structures under the label 'hybrids'.

The attribute categories (incentives, administrative controls, adaptive capacity and legal regime) proposed by Williamson (1991) as dimensions of governance structures, can be further elaborated to a much finer detail by incorporating insights produced by other relevant theories (e.g. agency and property rights) on

incentives, authority, bargaining and information, to name but a few. This line of thought can be pushed as far as arguing that the dimensions of governance are as many as the clauses of the most complex contract. Our only means of creating order and generality out of this complexity are the various available theories, such as those reviewed above.

Another reason why the transaction cost framework needs to be supplemented by more detailed constructs arises from a criticism often raised against transaction cost theory, namely that it is tautological. "*One could always invent a specification of transaction costs that could rationalise almost anything. By rationalising everything we might really be explaining nothing*" (Joskow 1993: 118). This is expected to happen when the notions we ascribe to our theoretical constructs are too broad and inclusive. For example, using the notion of transaction costs without discriminating between management, measurement, communication, bargaining and other costs, is one source of tautological accounts. Transaction cost theory highlights the significance of asset specificity, but this is not the only relevant dimension nor a sufficient explanatory condition. Goldberg (1989: 21-23) makes a similar criticism by highlighting the difference between transaction costs as the costs of the activities involved in transacting (e.g. negotiation) and transaction costs as the comparative efficiency of alternative governance forms. Although it can be argued that these two kinds of costs are not independent, the point remains that a closer examination of the nature and components of transaction costs is needed.

As different governance structures possess elements of markets and hierarchies in varying proportions, some are bound to be more akin to the ideal market, and some closer to the ideal hierarchy. All these arguments so far can be conceptualised with a continuum of governance structures as in Figure 2.2. The continuum is populated with generic categories of governance structures (e.g. joint ventures or franchising), which are considered to occupy narrower or broader ranges on the governance continuum between the two ideal extremes, the spot market and the pure hierarchy. That the pure spot market and hierarchy are fictional is explained

on page 94 above. In the diagram, internal organisation has been separated from the pure hierarchy in order to emphasise the fictional character of the latter and to show that internal organisation can be structured in a number of different ways, employing more or fewer elements of markets.

Actual governance structures as found in practice can be classified under a generic category but are expected to be fairly unique when the details of their composition are brought into consideration. Hence, they are expected to occupy discrete positions on the continuum, within the range of their category. For example, a given IT outsourcing contract may be classified under long-term procurement but when we examine its exact incentive structure, its adaptive mechanisms and its other elements of governance in detail, it is likely to be quite unique.

The governance continuum is derived from the theoretical arguments elaborated in the present section. Clearly, some of the debates in the IS literature that were reviewed in the previous chapter were also pertinent in shaping this concept. Fundamentally though, the governance continuum is a theoretical proposition drawn from a critical reappraisal of the economic theories reviewed in previous sections, especially transaction cost economics.

Comparative institutional assessment, in the sense advocated by Williamson, is applicable among governance categories. For example, time and materials contracting is relatively closer to spot market transactions as compared to inside contracting (Buttrick 1952), which is closer to the hierarchy. Comparative assessment is also applicable within the range of each category (i.e. between actual, unique, contracts), thus allowing us to populate these generic ranges on the governance continuum with observed governance structures.

Figure 2.3 demonstrates how the governance continuum might be interpreted in the context of IT contracts. Again, the governance categories listed are indicative and do not represent a one-to-one correspondence with the governance categories

listed in Figure 2.2. The internal IT department is not equivalent to a pure hierarchy to the extent that it may be decentralised, devolving decision and responsibility to end users and employing a chargeback procedure (see chapter 1, section 6).Contrary to some widely held beliefs, particularly in the IT literature, long term procurement of the kind employed in total outsourcing deals is very similar to internal organisation. In other words, internal organisation and such contracts employ very similar governance mechanisms (Ang and Beath 1993).

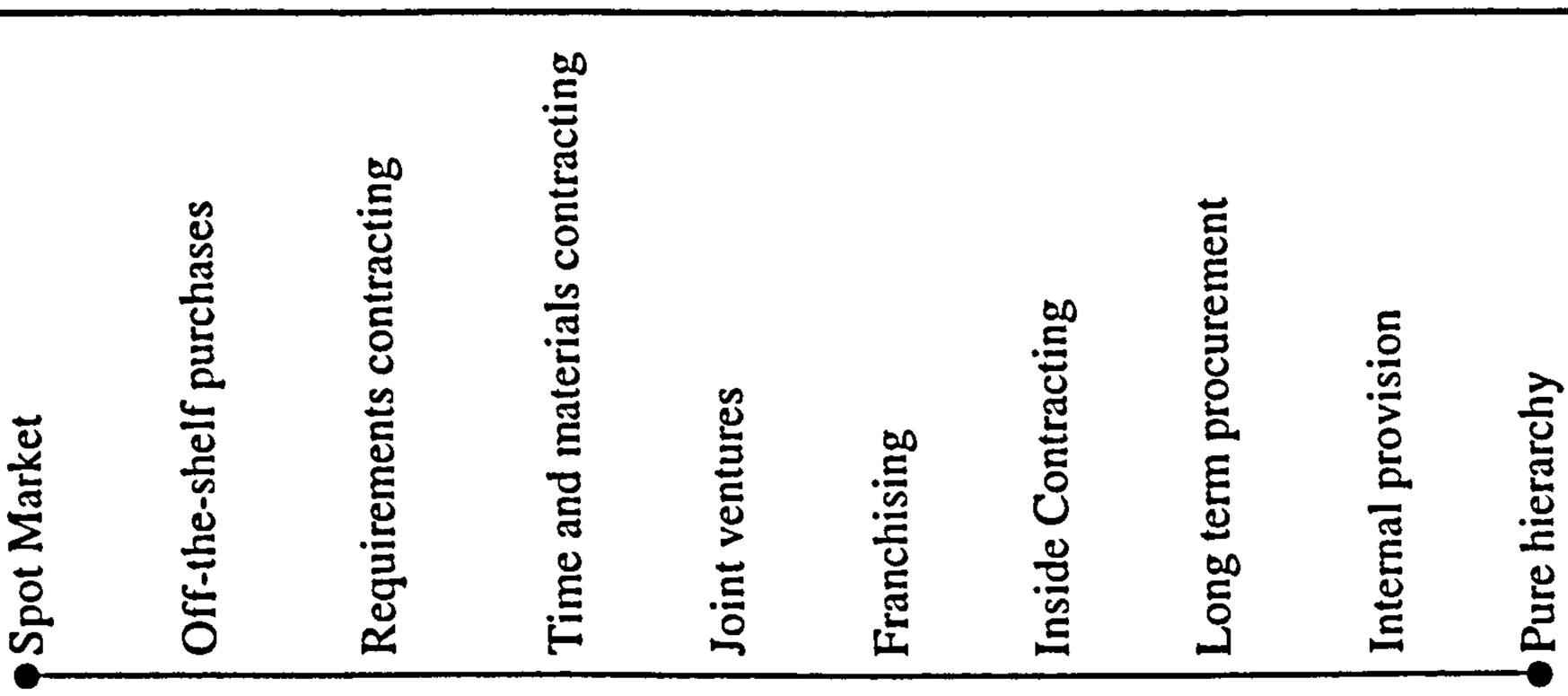


Figure 2.2

The Governance Continuum with indicative categories of governance

Between the two ideal types of the pure spot market and the pure hierarchy lies a broad range of categories of governance structures. These possess elements of both markets and hierarchies in varying degrees. They are placed closer to the spot market or the pure hierarchy according to a comparative assessment of these characteristics. Each category occupies a range on the governance continuum. Actual governance structures (contracts) are unique and can be discretely positioned under a category on the continuum.

Broadly, individual programming contractors (Figure 2.3) correspond to 'inside contracting' (Figure 2.2) and systems development contracts (Figure 2.3)are typically of the time and materials or the requirements contracting kind.Off-the-shelf software purchases are not pure spot market transactions because they are accompanied by warranties and possible maintenance agreements, which extend the exchange relation beyond the 'atomic' transaction.

Following the critique of transaction cost economics in the previous paragraphs, alternative governance structures are conceptualised as varying in the degree of certain attributes, rather than being different in kind. The attributes on which they vary are derived from the theories reviewed in earlier sections. These attributes

include, for example, asset specificity, incentive intensity, adaptive capacity, measurement costs, ownership allocation and reputation effects. Clearly, a more coherent classification of these attributes (or elements or mechanisms of governance)is needed. This is one of the contributions of the thesis, presented in chapter 5, section 4.

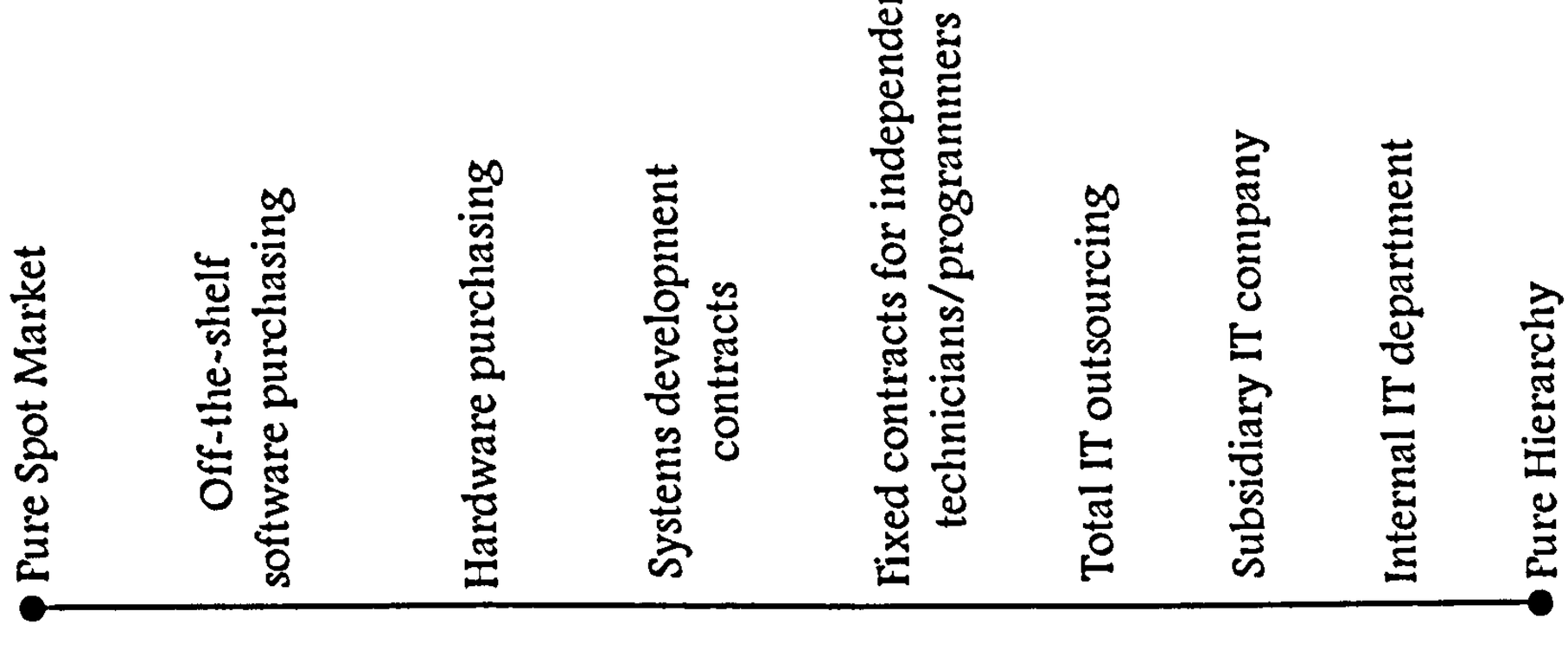


Figure 2.3

The Governance Continuum with indicative IT contracts

The microanalytic approach to economic organisation takes on two more dimensions, in addition to the above mentioned need to unpack the bundle of goods and services in a transaction and to consider their characteristics. First, the general governance alternatives (markets, hierarchies and hybrids) must be unpacked to reveal the many different governance possibilities. Second, the all-inclusive notion of transaction costs must also be unpacked to discriminate among several detailed elements of governance.

According to transaction cost theory, the purpose of economic organisation is to assign "*transactions (which differ in their attributes) to governance structures (the adaptive capacities and associated costs of which also differ) in a discriminating way*" (Williamson 1985: 18). The way this alignment is achieved in the context of the microanalytic strategy defined here, is by examining the characteristics of the goods or services being exchanged and customising the necessary elements of

governance. This is a bottom-up approach, as opposed to selecting from a 'menu' of alternative governance categories such as those listed in Figure 2.3

The governance continuum provides the platform on which the modified microanalytic strategy proposed here, the principle of comparative institutional assessment and the alignment of governance to transactions, advocated by transaction cost theory, make space for important contributions from agency theory, property rights, measurement issues, information and knowledge considerations and reputation effects.

Clearly, the role of the legal regime governing different types of governance structures, has been suppressed. Including it would have required a detailed account of its effects on efficient organisation, thus further complicating the task. However, inasmuch as the theories adopted embody elements of the law, the latter is not completely ignored. Moreover, by exogenising the legal regime it has been possible to alleviate (at least temporarily) the discrete character of internal organisation. It is the legal doctrine of forbearance that renders internal organisation discretely different to other governance structures (see section 2.2). Although certain governance mechanisms prevailing within hierarchies are unique (e.g. promotion as incentive and influence costs), analytical equivalents can be found in intermediate contracting structures. Therefore, the notion of a continuum from the pure spot market to the pure hierarchy can be sustained.

The first chapter concluded by identifying the fundamental research question as how to structure the efficient governance of the provision of IT systems and services. This chapter sets the theoretical basis for tackling this question in the rest of the thesis. The breadth and complexity of all the theories adopted is one of the problems addressed in chapter 5, by proposing an organising framework of the various ideas reviewed here. Meanwhile, chapter 3 explains the epistemological link between the IT outsourcing phenomenon, the theories of economic organisation and the framework delivered in chapter 5.

CHAPTER THREE

An Epistemological Basis for Research on the Governance of Information Technology Provision

1. Introduction

The purpose of this chapter is to explain the research approach adopted in this thesis by developing a meta-theoretical account of what is being attempted. This will also provide a basis for assessing what is being achieved. It is not simply a review of the advantages and disadvantages of the research methods employed. Rather, it is an effort to surface epistemological and methodological considerations that are often bypassed as implicit assumptions.

Arguably, every research project is able to make its theoretical claims only on account of a specification of epistemological and methodological considerations, even if unconsciously, implicitly or otherwise. On occasion, the nature and extent (and ultimately the value) of theoretical contributions may be self evident by virtue of the extent to which they appear realistic in explaining the phenomena they study. However, theoretical success is rarely so obvious mainly because research and theory development are conditioned by the social context in which they take place and, in turn, they shape the social objects of study. Some criteriaexternal to the theory concerned are, therefore, needed. However, what follows is not a complete review of alternative philosophies of science. This would be far beyond the scope and purpose of such a chapter within a thesis on the governance of information technology.

This thesis aspires to follow the principles of the philosophy of science typically recognised under the heading of transcendental realism. In the absence of an unequivocal unified philosophy of science, the adoption of a particular philosophical position is necessarily subject to all those personal and social conditions that frame human thought and action. The appeal of transcendental realism, which led to its adoption in this project, rests mainly on its adherence to an independent reality and its integrative accommodation of human subjectivity, interpretation and agency alongside this reality. Moreover, realism makes no *a priori* judgements for or against particular methods of research; their contribution to a research project depends on the context and process of their use.

Given this philosophical position, the purpose of this chapter can be more specifically redefined as aiming to assess the nature of this project and the challenges it confronts, including the empirical methods it employs, if it is to conform with this philosophical principle.

The rest of the chapter starts with an exploration of methodological and philosophical discussions common in recent IS literature, with a view to showing the importance of making ontological and epistemological assumptions explicit when contemplating alternative research methods. A brief outline of the transcendental realist ontology follows. Next, the nature and aims of the project presented in this thesis are restated in the context of this philosophy. Mainstream economics has been criticised for its insistence on positivist research. The main tenets of this critique and the possibility of incorporating extant economic theories in a realist theory is the subject of section 5. Section 6 examines the methodological implications of realism for this research project and explains the exact methods employed. The concluding section reasserts the role and significance of this chapter in the flow of the thesis.

2. Epistemological and methodological concerns in IS research

In the area of information systems research, lengthy discussions have risen as to the methods, methodologies and approaches that are appropriate for the questions encountered by IS researchers (e.g. Fitzgerald et al 1985, Cash and Lawrence 1989, Orlikowski and Baroudi 1991, Galliers 1992, Mingers and Stowell 1997). One distinction commonly made is between quantitative and qualitative research approaches (Kaplan and Duchon 1988, Galliers 1991, Myers 1997). Literally, the distinction is between using quantitative measures and using qualitative observations. Logically, this distinction is extended to differentiate between the methods of data collection, analysis and inference that are appropriate to each kind of empirical evidence. The debate often revolves around two apparent deficiencies of the quantitative approaches (e.g. Klein and Lyytinen 1985, Pervan and Klass 1992). First, quantitative approaches are criticised for over-simplifying the complex, emergent inter-relationships between the objects of our study (information systems, people, organisations etc.) and reducing them to linear, sequential associations of cause and effect. Second, by starting with preconceived hypotheses, it is argued, quantitative approaches ignore the context in individual circumstances and supplant it with some sort of 'scientific objectivity'.

In contrast, qualitative approaches present the researcher with the opportunity to reveal the complexity and appreciate the context of the social setting being studied. The assumption here is that elements of the context, most notably the subjective and inter-subjective meanings ascribed by the members of a given social situation to the objects of study, cannot be easily measured and that complex relationships are not easily represented (and analysed) in mathematical form. Therefore, quantitative approaches are not adequate to cope and qualitative approaches are called for.

Moreover, if research is to inform the practice of developing and managing information systems, it needs to be able to take into account the complexity of the

'real world' within its context. Inasmuch as qualitative research is better at contextualising the research process and outcomes, it is argued, it emphasises relevance, whereas quantitative approaches, which favour experimental controls and mathematical sophistication, emphasise scientific rigour. The debate often turns to a dichotomy between relevance and rigour (e.g. Galliers 1995). However, as discussed further on, there are good reasons why the two are not mutually exclusive.

One suggestion hidden in the above argument is that some settings of interest to research (or some aspects of a setting) lend themselves to quantitative or qualitative approaches, to a greater or lesser extent. For example, with enough perseverance we might be able to measure certain attitudes or meanings, as it actually has been the case. In addition, when we seek generalisable patterns of association quantitative methods are more appropriate whereas when we are interested in deep contextual understanding qualitative methods are better suited. Consequently, it has been argued that certain research problems are better approached quantitatively and others qualitatively.

A second suggestion hidden in the above argument is that qualitative and quantitative approaches are better suited to different stages of the research process. If qualitative approaches are good for revealing complex structures and context-specific definitions (meanings), then they are more appropriate to the early, exploratory stages of research. These structures and definitions can inform quantitative methods (as the basis for formulating hypotheses), which, in turn, appear to be better at establishing the existence of causal relations with a high degree of external validity. Therefore, quantitative methods, it is argued, are more suitable for theory development and testing. Theories thus developed can also be subsequently evaluated using a qualitative approach. The rationale is that such an in-depth study is a good means for re-examining the reliability and validity of formal theories, mainly for purposes of falsification (Yin 1989).

The argument crudely sketched here underpins the taxonomy of research approaches proposed by Galliers (1991). He concentrates on research *approaches* as generic '*ways of going about one's research*' (p. 329) and proposes 14 categories including, for example, field experiments, surveys, case studies and action research. In a way similar to the points made above, Galliers (1991) argues that researchers should choose suitable research approaches according to the object of study (e.g. individual, organisations or society) and the stage of their research process (e.g. theory building, testing or extension).

Following the same line of thought, the case is made for using a pluralism of methods and approaches in combination, within the same research project (Kaplan and Duchon 1988, Orlikowski and Baroudi 1991, Galliers 1995, Silverman 1998). The main advantage is that sources of evidence can be triangulated and that more light can be shed on and, hence, more insights can be gained from the same evidence.

The debate on research approaches is often confined to substantive matters pertaining to the objects and theories of our research (e.g. whether they lend themselves to measurement or whether subjective meanings are deemed significant for the purposes of a given research question). However, inevitably such discussions rely on explicit or implicit assumptions regarding the locus and structure of reality and the nature of scientific knowledge. These issues are dealt with at a higher level of abstraction, namely at the levels of ontology and epistemology respectively. A variety of arguments and propositions have been advocated by different philosophies of (social) science. It is worth touching upon some key issues and relating them to the philosophical position underpinning this thesis.

A philosophical dichotomy often encountered in the IS literature is between (logical) positivism and interpretivism[1]. The distinction is sometimes wrongly equated with the use of different methods, quantitative and qualitative respectively. It just happens that positivism tends to favour quantitative research methods whereas interpretivism tends to favour qualitative methods. However, this correspondence is circumstantial and not necessary. Indeed, qualitative methods and evidence are easily accommodated within a positivist research programme. In principle, research methods and techniques are independent of epistemologies and ontologies (Myers 1997). For example, case study research can be positivist (Yin 1989) or interpretive (Walsham 1993) and so can action research (Clark 1972, Elden and Chisholm 1993). Moreover, in principle nothing precludes quantitative techniques from being constructively employed within an interpretive research programme.

The difference between these two philosophical positions lies in their different conceptions of reality and of scientific (as opposed to lay) knowledge. For positivism, real is whatever can be observed and recorded in an independent (objective) way. The purpose of science is to establish causal links between observations. Simply put, a description of such causal relationships is theory. Since reality is identified with observed experience, causal relationships are to be found in constant conjunctions of observations. In broad terms, this is the foundation of statistical significance. For example, if we can observe enough workers whose productivity has improved after computerisation (and assuming we adhere to sampling and other statistical requirements), then we can deduce that computerisation is a cause of improved worker productivity.

Interpretivism as ontology posits that in the social realm, reality is not objective but it is embedded in the subjective meanings that human agents ascribe to objects of

---

[1] Each of these two terms is used to refer to a family of similar variants of a philosophical direction (and a variety of related terms) but a discussion of such refinements is out of the scope of this thesis.

their environment. Individual predespositions and inter-subjective meanings constantly perpetuate and change each other. Interpretive research confronts social settings in their own terms and aims to develop theoretical explanations (interpretations) based on the definitions and meanings held by the members of the social setting itself. Interpretivism as epistemology posits that our knowledge of reality is a social construction. Therefore, objective and value-free evidence and theory are not attainable. As a consequence, weak constructivism, a strand of interpretivism, posits that the beliefs and meanings of the participants being studied can be accessed and documented in their original, pure form. In this sense interpretivism can complement positivist research by informing hypothesis formulation and interpretation of results. Strong constructivism, another strand of interpretivism, argues that in designing and carrying out social research, the researcher essentially reconstructs the reality being studied by interjecting his or her own predispositions. This version of interpretivism is incommensurable with positivism.

For the sake of compliance with recent discussions in the IS literature, it is worth reporting a philosophical position that is gaining increasing acceptance by IS researchers, namely critical theory (Orlikowski and Baroudi 1991). Critical theory adopts some fundamental tenets of interpretivism but makes significant further claims. Ontologically, social reality is systemic in the sense that objects exist and are meaningful only within a broader web of inter-relationships with other objects. Moreover, these social systems are historically constituted and defined in a continuous change process from the past, via the present and into the future. Change is triggered by contradictions and conflicts inherent within social systems. Emphasis is placed on how things are and importantly, on how things can be (potentiality). At any moment in time, human agents are entrenched in systems of beliefs, which, in conjunction with the material conditions of social reality, constrain their understanding of their historical reality and impede their further historical development and attainment of their potentiality. Epistemologically, the purpose of science is to make human agents conscious of these limitations, of

existing contradictions and of the potential directions of change. Thus theories adhering to this philosophy are emancipatory. In this sense, the claim of interpretivism that science must reveal the subjective beliefs of human agents within their social context is not enough; science must transcend those entrenched beliefs and at least enable human agents to decide upon and take action to transform their social reality (if not prescribe action on behalf of human agents).

This brief exposition of three oft-quoted philosophies of social science serves to emphasise the need to refer to a meta-theory when selecting research methods, developing theoretical arguments and evaluating research results. When alternative research approaches are debated within the bounds of a research question or theoretical perspective and without reference to any external criteria, any approach can be justified on its own terms. In practice, the choice of a research approach is likely to be influenced more by the dominant paradigm within the research community working on this subject area than by any inherent merits of either approach (Orlikowski and Baroudi 1991, Walsham 1995). To any suggestion of a proponent of qualitative methods that the latter are better at capturing elements of context and subjective beliefs, the proponent of quantitative methods can counter-argue that with enough perseverance sophisticated measurement instruments can be developed. To any suggestion of a proponent of quantitative methods that modern techniques of data analysis can manipulate complex inter-relationships and replicate experimental controls, the proponent of qualitative methods can always counter-argue that social relations are even more complex.

Moreover, as mentioned previously, the same methods can be successfully employed under different epistemological positions. Case studies can be carried out in different ways in a positivist as well as in an interpretivist research programme (Benbasat et al 1987, Yin 1989, Cavaye 1996, Myers 1997). Similarly, participant observation can be employed to record observed practices in a detached manner for the purposes of (qualitative) hypothesis testing (Silverman 1998) or they can be

employed in a more engaged manner to reveal and debate hidden assumptions, meanings and beliefs (Pettigrew 1985). Experiments and quantitative techniques of data analysis can also serve the purposes of an interpretive or critical research programme, even if they are not sufficient on their own.

How can relevance and rigour be assessed, without a conception of the structure of social reality that we want to explain in a relevant way and without a conception of what constitutes valid and rigorous knowledge? Statistical techniques define specific tests of validity, which cannot be literally replicated in qualitative methods. However, theoretical validity can be argued by other means, notably by reference to an ontological and epistemological position (Lyytinen 1987, Tsoukas 1989, Numagami 1998). An argumentative or a grounded interpretive theory can be as rigorous as a mathematical theory but the criteria for achieving and assessing rigour must be external to the theory concerned, stemming from an epistemological position.

Whether such philosophical assumptions are explicitly articulated or not, any research design or any argument about a research design (approaches, methods, techniques) is bound to rely on hidden philosophical assumptions (and this is an interpretive statement in itself). Therefore, the upshot of this section is that it is useful and desirable that a researcher makes his or her ontological and epistemological assumptions explicit. These assumptions guide the design of research, the choice of methods, the development of theory and the evaluation of the whole project by peers, clients and other third parties. This brief review of the main philosophical and methodological arguments within the IS literature also serves as the backdrop for the philosophical and methodological position adopted in this thesis and explained in subsequent sections.

One does not have to be an expert in all the philosophies of science and their often esoteric nuances in order to articulate the meta-theoretical assumptions permeating his or her own research project. Admittedly, opening such a philosophical

discussion sometimes amounts to opening Pandora's box. There are many different theoretical positions and even more refinements and variations within each school of thought. Understanding all the arguments and counter-arguments and adopting an informed opinion is a job for experts. It is possible, however, for any researcher in any field of social enquiry to educate him or herself in the basic ontological and epistemological questions, in order to be able to make explicit his or her assumptions when devising a research design.

Another problem is which side to take in the philosophical debates. There does not seem to be a further meta-level of theory from which to draw criteria for selection. Although some philosophies of science present themselves in an exclusionary way, it is possible to argue that none of them holds the absolute truth. In the absence of external criteria (except the rules of logic) and of an ultimate truth (except, perhaps, faith), adoption of one or another philosophical position turns either to a utilitarian criterion (whether a given philosophy serves our purposes and hidden agendas) or to the researcher's cognitive and intellectual predispositions (which, in turn, may be socially and historically constituted) or both. The realist philosophy of science, which is adopted in this thesis and briefly presented in the next section, adheres to an independent reality and accommodates subjectivity, interpretation and human agency in an integrative way, alongside this reality.

3. The main tenets and implications of transcendental realism

A complete account of the realist philosophy of science is beyond the aims of this chapter. Instead, a brief outline of this perspective is provided here, based mainly on the works of Bhaskar (1978, 1979), Outhwaite (1983, 1987), some of the contributions of Chalmers (1982) and its implications for economics elaborated by Lawson (1997). By setting the scene in a concise manner, subsequent sections can highlight issues of more direct relevance to this research.

The realist philosophy of science turns its attention to ontology by asking what should the world be like if scientific enquiry and knowledge are to be possible. Implications are consequently drawn for epistemology which concerns itself with the nature of scientific knowledge and its explanatory capacity. Methodology refers to the practical means of acquiring or developing such knowledge. As it turns out, the realist philosophy informs the capabilities and limitations of different methodological approaches but does not prescribe a specific programme.

For the realist philosophy, the world comprises three domains, namely the real, the actual and the empirical (Table 3.1). In brief, we acquire experiences from the empirical domain. These experiences are subjective representations of events that occur in the actual domain. Actual events are, in turn, contingent products of the interaction of objects of the real domain. The three domains are irreducible to one another.

	Real Domain	Actual Domain	Empirical Domain
Objects	✔		
Events	✔	✔	
Experiences	✔	✔	✔

Table 3.1

The three domains of the world according to transcendental realism
The checkmarks (✔) indicate the domain of reality in which objects, events and experiences reside.
Amended from Tsoukas (1989).

The real domain is populated by objects with inherent structures and causal powers. Different objects in the real domain may be related internally or externally. Two objects are internally or necessarily related when the existence of each depends on the existence of the other (e.g. employer and employee or information system and user). Two objects that are not internally related can, potentially, be externally or contingently related when they happen to interact (e.g. a married couple and a court of justice or an information system and a hacker).

When the causal powers of objects are triggered the respective mechanisms are set into operation which, in turn, generate some effects. These mechanisms are not necessarily in constant operation. In fact, they may exist but never actually be triggered in operation to produce actual events. Therefore the real and the actual domains are separate. Moreover, a mechanism may be in operation but not necessarily be observable. Therefore, the empirical domain is also separate from the other two. Gravity is the typical example here: it is a mechanism constantly in operation but we do not constantly observe its effects, namely objects falling to the ground. Another example, related to the subject of this thesis, is the power of contractual clauses and the legal system to enforce promises and/or damages. This is a mechanism that is constantly there but it is rarely triggered since contracting parties usually prefer alternative (less costly) means for resolving their disputes (Macaulay 1989).

The events that are actually (i.e. in the actual domain) produced are contingent on a variety of (typically unpredictable) conditions that may emerge in the interaction of objects and of their causal powers. The events produced are not necessary consequences of the mere existence of real objects, but contingent results of their interaction. In other words, the actual domain is an open system where real objects interact and events are produced.

The ontological separation between the real and the actual domains can be illustrated with an example from meteorology. Specialists have a fairly complete knowledge of the structure and causal powers of 'objects' in the atmosphere and can provide detailed and precise explanations of the phenomena that occur. However, weather forecasts are notoriously inaccurate. This is not because the knowledge of the *real* causal mechanisms is incomplete (since *ex post* explanations are feasible and fairly complete) but because the *actual* weather conditions that emerge take place in an open system and as such they are inherently dynamic and unpredictable.

The empirical domain is the world of perceptions and experiences of human agents, including scientists and lay people alike. Events occur independently of their identification. Some events may not be experienced at all and the perceptions that different individuals hold of those events that they perceive are conditioned by a variety of factors including their senses, their instruments (such as microscopes or questionnaires), their subjective preconceptions and their social environment.

For example, wars can be attributed to some sort of inherent tendency of groups of people to pursue or protect their collective interests using military force. In a simplistic way, this would be a mechanism in the real domain. A particular war is a result of various historical contingencies occurring in the actual domain. The same war is experienced (in the empirical domain) by different participants and observers (e.g. winners, losers and mediators) in many different ways.

The term 'transcendental' that features in the name of this philosophy refers precisely to the fact that the objective reality advocated resides beyond our experience of it and beyond its manifestation in actual events. In other words, reality transcends both our knowledge and experience of it as well as the events produced by it. For example, electricity exists in nature even when no electrical phenomena (events in the actual domain) take place (such as a lightning or an electric lamp being lit) and it has always existed even before we understood it as such (knowledge and experience in the empirical domain). As a social example, most would agree that the information revolution is a historical change comparable in magnitude to the industrial revolution (Angell 1998, Castells 1996). Though we are all living through this change, we are not constantly aware of it nor do we understand it fully (knowledge of the real mechanisms). Moreover, this revolution is manifested in a number of technological, economic, political and social ways (events in the actual domain) but not all of us experience all these events and not in the same manner (observations in the empirical domain).

Once it is recognised that the objects of science belong to an objective reality, independent of our experience and knowledge of it, the purpose of scientific theory is to postulate an account of the structures, causal powers and mechanisms that can produce the phenomena under investigation. Scientific enquiry, therefore, needs to transcend both the empirical and the actual domain in order to reach their intransient generative mechanisms. This is not to say that experiences and events are not real and do not merit investigation. On the contrary, the explanation of a particular event or experience will also require an understanding of the contingencies that lead to the particular state of affairs and the parameters that condition our perception of it. As indicated in Table 3.1, contingent events and experiences are part of reality. However, scientific enquiry is not exhausted at these two transient levels, without an adequate theory of the intransient mechanisms that produce them. For example, it is not sufficient to record the behaviour of erupting volcanoes. Both the event of eruption and our observations of it are contingent (circumstantial). The causes lie in the structure of the earth, which we do not observe directly. Similarly, the seasonality of consumer demand for turkeys in some countries is not sufficiently explained by the coincidence of certain holidays. Such an account would collapse once the holidays are maintained while the respective traditions are not upheld (or *vice versa*). The historical antecedents, the meanings and the social structures reproducing these traditions need to be explained in their own right.

A further property of the real domain is stratification. Objects can be conceived as made up of component parts and as being themselves components of larger structures. In each case, the whole has emergent structure and causal powers and is irreducible to the sum of its parts. For instance, a firm comprises a set of human and physical resources but it is much more than their sum. Similarly, water is used to extinguish fire although both oxygen and hydrogen are highly flammable. Scientific enquiry is therefore never ending. Once an acceptable (perhaps temporarily) theory of some mechanisms has been attained, the component and containing objects will have to be further explored. For example, the power of

water to extinguish fire may be an acceptable first theory. Further research uncovers how oxygen and hydrogen combine to such an effect.

One reason why we need to accept the existence of the intransient real domain is that it renders scientific enquiry possible. If reality was exhausted with contingent events or experiences then unique phenomena would remain unintelligible. Therefore, the objects of reality need to be enduring or at least enduring enough for their scientific discovery to be possible. It follows, that scientific knowledge is transient, at least relative to its objects of study.

To place the realist perspective in context, it is useful to contrast it with some other extant approaches (see also the discussion in section 2). The positivist tradition identifies laws and necessity with empirical regularities. In terms of the realist framework, it collapses the real and the actual domain on to the empirical by recognising regular contingent coincidences as real. Positivism relies on deductive reasoning. For example, to explain a temporary increase in the demand for turkeys, the positivist needs a set of initial conditions (that it is the period before Christmas) and a covering law (that demand for turkeys is higher before Christmas). The conventionalist strand of positivism accepts a reality independent of our knowledge of it but posits that explanation can be based on purpose-built and provisional assumptions. The critique of theory is, therefore, suppressed. In a similar vein, for pragmatism all that exists is ad-hoc theories. Some strands of interpretivism identify as real only the meanings that individuals ascribe to experience. It is clear that all these approaches collapse ontology to experience and entail the risk of extreme relativism, that 'anything goes'. Kant's transcendental idealism finds necessity in the theories of science, only in this case these theories are intransient. This approach disregards the social construction and conditioning of science and hence its fallibility and transience. Critical theory comes closer to transcendental realism. It emphasises the objective reality of the material conditions of society but realism makes even stronger ontological claims. The importance placed on social critique is shared by both theories.

Even in this short exposition, the validity of realism for the natural sciences appears quite intuitive. As it happens, the development of realism by Bhaskar started with an assessment of experimental practices in the natural sciences. In order to adopt realism in the social realm, several issues need to be contemplated, mainly reflexive and intentional human agency and human choice.

In the social realm, transcendental realism can be challenged by an apparently simple question: how is it possible for real social structures to exist independently of their manifestations and of our experience of them, if human agents create, reproduce and change them? This question raises the classic paradox between social structure and human action (Poole and Van de Ven 1989). If social structures did not exist in their own right and were reducible to human agency, then the human agency would not be possible, at least not as we know it. However, social systems, rules, relations, traditions and the like precede any individual and condition his or her actions. If, in turn, human agency was to be denied and identified with social structures, then there would be no space for human choice, i.e. the capacity of individuals to act otherwise. It follows that human agents and social structures are internally related in the sense defined previously. Neither is identifiable with the other, yet each is necessary for the other's existence. If we accept that human agents can take purposeful and intentional action, can reflect and act upon their knowledge (whether scientific or lay) and can choose their actions, then social structures are susceptible to change and reproduction as a result. In other words, human agents are active members of the real domain and have the capacity to reproduce and/or transform it.

It can be shown that despite these differences between the natural and the social world, transcendental realism can still be sustained in the latter. Intentional choice and action presuppose some form of knowledge which, in turn, presupposes the objects of knowledge are enduring enough to facilitate their coming to be known. Also, in reverse, the development of a social scientific theory will affect the

decisions, choices and actions of those individuals who get to know it. They will adjust their behaviours thus changing and/or reproducing the social structures. A new theory might then be needed to explain the changed structures and behaviours. This relative intransience of social objects over sufficient time-space windows sustains and is sustainable under transcendental realism. Moreover, the realist ontology sustains those crucial elements of the social world, namely intentional human choice and agency. In contrast, the positivist strands deny the possibility of choice since human actions have to follow the empirically deducted 'laws' of necessity (Lawson 1997).

Predictive power has traditionally been held as the acid test of any scientific theory. Given the realist ontology, it can be shown that prediction is ultimately impossible, not essential and, in the social sciences, undesirable. Given the contingent nature of the actual and the empirical domains, successful prediction is possible only in the closed system conditions of experiments. Experimental conditions are, however, very rare in the social world, despite the effort made in this direction by modern techniques of data analysis. Such techniques attempt to replicate conditions of closure by incorporating more variables of the phenomenon of interest and by endogenising as many situational variables as possible. Ultimately, however, conditions of experimental closure are almost impossible to achieve in the social realm. In addition, forecasts in the social realm can be self-fulfilling or self-negating by virtue of the reflexive nature of human agency to react to such knowledge (Numagami 1998).

Prediction and forecasting are, also, not essential for policy formulation and implementation. All that is required, indeed, is an understanding of the real underlying mechanism according to which society acts, reacts and changes within given time-space frames. Actual events are always contingent and unpredictable. Finally, prediction in the social realm is undesirable because it leads to the possibility of engineering and controlling human activity. The latter, in turn, leaves no space for human freedom and choice.

At the level of methodology, realism has practically no firm directions to give. Substantive theories are developed in an incremental way and are assessed in comparative terms. In particular, for each phenomenon of interest, a theoretical account will have to be compatible with the realist ontology and postulate real causal mechanisms as well as possible contingent factors that explain the occurrence of the specific empirical events being studied. This account will have to start from the (lay) interpretations that human agents ascribe to the structures and phenomena of interest within the context of their everyday activity. These subjective interpretations constitute part of the reality of the phenomenon but do not exhaust the causal mechanisms that produce it. Additionally, each postulated theoretical account will have to make reference to other extant theories of relevance.

Scientific knowledge is thus socially conditioned. The explanatory power of a theory will have to be assessed according to its capacity to provide sufficient explanations of the phenomena of interest. Theories are also always provisional, in that they are often compared and criticised against competing or complementary explanations. Complementary explanations may refer to related phenomena or to further levels of (the stratified) reality. Given the stratified structure of reality, multiple alternative theories can be sustained unless, of course, a better theory explicitly rejects earlier explanations. Different aspects or different levels of abstraction of reality can be explained by different theories, each of which can be possibly judged as adequate in its own right. The discussion of paradox in social theories by Poole and Van de Ven (1989) succinctly demonstrates this possibility. For example, Newton's theory is still valid although it has been superseded by Einstein's theory. And the latter can stand alongside quantum mechanics although, at present, they are incommensurable. While theories are transient, i.e. subject to revision or adaptation, consistency with the realist ontology and adherence to the principle of critique ensures that science is not reduced to extreme relativism.

4. The project undertaken in this thesis

This thesis is motivated by the widespread occurrence of large outsourcing agreements for the provision of information technology services since the late 1980s. A substantial amount of research that ensued has indicated various causes, reasons, and factors of success or failure of such practices (chapter 1). As elaborated in the ends of chapters 1 and 2, this thesis focuses on the governance structure of such arrangements.

In particular, the governance structure of the provision of IT services is postulated as a general mechanism determining, to a non-trivial extent, the organisational and contractual forms we observe and their success or failure. Clearly, the causal powers of the governance structure are not the only determinants of the respective practices that are actually implemented and of their antecedents and consequences. Other research has elaborated on the role of intra-firm politics, strategy and the bandwagon effect. This thesis abstracts from such other parameters to study governance structures in more detail. Empirical findings presented in chapters 4 and 6 help re-position the abstraction into its broader context and locate its significance.

In order to illuminate the operation of governance structure as a generative mechanism, this thesis draws heavily on extant economic theory. The task is synthetic, in that it attempts to bring together various contributions from a wide range of relevant, albeit in many respects different, theories (chapter 2). The goal is to produce a theoretical account that is both valid theory and a useful guide for decision making in practice.

The resultant framework (chapter 5) is presented as primarily a normative guide for decision makers, although it consists of positive theories of governance. It consolidates a coherent account of what a governance structure is, what it consists of and how it operates. It postulates in an organised manner a number of partial

mechanisms that in conjunction give shape to the governance of IT service provision. Not all of these mechanisms are expected to be necessarily operative or of the same significance in every actual instance. They are however potentially operative and from a normative point of view they should be taken into account whenever relevant.

While the framework is presented as normative, it is also (necessarily) an explanatory theory. The majority of its substantive content is already available, albeit dispersed, in the economics literature. Its original explanatory contribution lies in the synthesis and repositioning of extant theory (see end of chapter 2) and in its specific formulation to address the peculiarities of IT governance[2]. Moreover, the production of a theory directly usable in practice has been one of its original objectives.

Briefly, the realist objective has been pursued by repositioning the positive economic theories of governance as generative mechanisms having the capacity to produce the postulated outcomes. The latter are, therefore, potentialities rather than predictions. The normative objective has been pursued by studying the structures and practices pertaining to IT outsourcing *in situ.*

As also explained in section 6.2, despite the normative form of the proposed framework, a contingency structure has been intentionally avoided. Since, according to realism, actual events and situation-specific contingencies are varied and unpredictable, a contingency structure, (apart from futile) would be too constraining and less informative in practice. By explaining some mechanisms of governance that are potentially at work, the proposed framework is intended to promote criticism, appraisal and debate among practitioners regarding the best course of action, rather than to offer pre-fabricated solutions which do not relate to

---

[2] The specificity of the framework to IT versus its general applicability to other business activities is discussed, among other issues, in chapters 1 and 7.

the actual context. This is consistent with the principle of social critique and the argument against prediction that realism advocates (and, to a significant extent, critical theory shares). Numagami (1998) succinctly demonstrates the point: "*Given that human beings have reflective capacity, the realistic goal for management researchers is not to teach scientifically naive practitioners law-like regularities confirmed by a scientifically rigorous method, but to encourage reflective dialogue among researchers and practitioners to generate more reasoned discourse on today's key management issues.*" (p. 4).

To demonstrate the significance of the framework's theoretical contribution and practical applicability, two issues warrant further discussion. First, economic theory, including the strands reviewed in chapter 2 and subsequently employed, have been criticised as typical cases of misconceived positivist research. Their incorporation into the realist ontological position adopted here requires important qualifications. Second, the empirical methods employed determine to a significant extent the success of the project, both in terms of achieving its realist objectives and in terms of ensuring its practical (normative) relevance. The instrumental role of the case studies therefore requires further elaboration. These two issues are addressed in turn, in the following sections.

5. Incorporating economics into a realist theory

Economics, as a discipline, has been widely criticised for its failure to produce realistic theories, to inform policy making and to predict economic development (see for example, Blaug 1980). Much of the economics literature consists of formal models of high technical sophistication which achieve little more than elegant exercises in mathematics with more internal consistency than practical relevance (Jensen 1983). Empirical work has, at large, failed to falsify any theory unambiguously, as the positivist tradition of the mainstream dictates. It also appears that any theory (and often contradicting theories) can find confirmatory evidence in an appropriately manipulated set of data. The practice of empirical economics

seems to have distanced itself from the principle of falsification in favour of confirmatory exercises which are more easily attainable (Lawson, 1997).

Part of the blame is ascribed to the overly simplistic and typically fictitious assumptions imposed on theoretical models. Such assumptions include the omnisciently rational nature of economic agents, the profit maximising purpose of firms or the utility maximising behaviour of consumers. In response, several departures from the orthodox economic programme have taken on the task of correcting the failures of the mainstream by relaxing or otherwise adjusting these assumptions. Transaction cost economics is one such example. As it turns out, however, these alternatives still face significant problems in empirical research. No matter what claims can be made regarding the superiority of their assumptions, they often lack the technical sophistication of the mainstream but share its empirical failure. Despite a growing amount of empirical studies in these areas, they face additional problems in formulating empirical tests comparable to the established mainstream practices mainly due to the difficulty of measuring key concepts (such as asset specificity or bounded rationality).

Another response has been to pursue even more sophisticated mathematical techniques in theoretical modelling and data analysis methods for empirical work, with a view to improving the explanatory power of the economist's tools. While tools and techniques are being improved and developed, successes in substantive theory have not ensued. The academic community in economics appears to be caught in a vicious circle of more technical sophistication, arbitrary assumptions and theoretical propositions of little relevance to the actualities of the economy.

Transcendental realism can help shed new light on this criticism (Lawson 1997). From a realist perspective, the problems of mainstream economics are due to its (sometimes implicit) insistence on a positivist research programme and a deductive mode of reasoning. In econometrics, this is clearly the case. Scientific laws are sought in constant conjunctions of events. The explanation of economic

phenomena is subsequently deductive, given a set of initial and boundary conditions and the relevant empirical laws. Therefore, from the realist point of view, the relative failure of econometrics and its apparent capacity to find corroborative evidence for different, often opposing, theories, is no surprise, since empirical regularities (in the empirical domain) are coincidental rather than necessary (as are the structures of the real domain). The improvements sought by econometricians are in the direction of measuring more aspects of the phenomenon of interest and by endogenising more situational variables that are seen to impinge on that phenomenon. Such efforts attempt (often unconsciously) to secure conditions of experimental closed systems, akin to the experiments conducted in some natural sciences. In the realist terminology, these attempts seek to establish conditions of closure that minimise the contingent interrelationships of the actual and empirical domains in order to reveal the direct effects of the real mechanisms in operation. However, the openness and complexity of actual economic systems are pervasive.

Theoretical modelling in economics is also based on the same positivist-deductivist conception of science. Axioms and assumptions are used as empirical actualities in the place of empirical laws. Often, they are selected and configured so as to suit the tractability of the mathematical calculus, which is used to deliver various solution concepts. These solution concepts are also postulated as assumed features of actuality. Mathematical sophistication is sought to accommodate more realistic assumptions. Whether such models are empirically tested or not, it is implicitly held that an appropriate research design could be formulated for an empirical test aiming to confirm or falsify the results of the model on the basis of the existence or absence of corroborating empirical regularities. However, given the typically superficial nature of the assumptions held and the open and contingent nature of actual economic systems, it is not surprising that empirical support for formal models is scarce.

The theories reviewed in chapter 2, including transaction cost economics, are not exempt from this criticism. Although, as their proponents claim, they constitute a departure from orthodox economics, the difference is limited to the substantive assumptions and propositions. The underlying conception of science remains positivist. The question then begged is whether and how such theories can be incorporated in a realist theory that this thesis aspires to develop.

That economics is based on a misconceived ontology does not necessarily mean that all the economic theory thus developed over nearly two centuries has to be discarded. Transcendental realism as a philosophical position cannot decide on substantive matters (i.e. which one is a good theory and which is not). Also it does not, at the outset, rule out any empirical methods. It does, however, have implications for assessing the nature of theoretical claims. Moreover, nothing precludes that a positivist research programme cannot by chance, by insight or otherwise produce valid theoretical explanations. However, realism requires that such explanations be postulated as potentialities rather than actualities.

For example, it cannot be denied that economic agents possess some kind of rationality. The manifestation of rationality in practice, though, is conditioned by a number of parameters pertaining to the decision-making context and the individuals' personality. To study human rationality experimentally would require controlling all aspects of one's personality and all situational conditions. While this is ultimately impossible in practice, it is possible to develop explanatory accounts of observed phenomena that separate the real mechanisms of rationality from the contingencies that lead to decisions and actions and from the factors that affect the observer's perception of ensuing events. Therefore, the observation of apparently irrational actions does not by itself falsify the assumption that human agents are rational. An examination of the context and of the individuals' private preferences may reveal that what appears rational to an observer with different preferences and social conditioning is actually perfectly rational for the individuals being

studied (see a related criticism against transaction cost economics raised by Lacity and Willcocks 1995).

This is the way in which the theories of chapter 2 are employed in subsequent chapters. These theories are conceived as possible mechanisms that come into operation in observed governance structures, although their relative importance and manifestation will vary across different circumstances. Moreover, they are not considered to constitute a complete account of all mechanisms that can determine actual governance structures. However, they do cast a broad net over several different possible mechanisms.

Awareness of these mechanisms and of their potential effects provides the basis for a constructive dialogue among researchers and/or practitioners on a course of action that is appropriate to the specific context of the situation. For example, the theory postulates various incentive effects of asset ownership, of reward structure and of performance target setting over time. The extent to which these mechanisms are efficacious, significant or even relevant in any particular contracting case will vary from case to case and, therefore, different contractual and organisational arrangements will be made. This is the way economic theories are used in this thesis and incorporated in the proposed framework in chapter 5.

The theoretical basis for this thesis laid out in the very last section of chapter 2 certainly conforms with this principle. There it is argued that the idealised, fictitious concepts of markets, hierarchies and hybrids are unrealistic and a continuum of actual governance structures is proposed instead. Each of the latter, it is argued, is determined contingently by a range of possible mechanisms, such as incentive schemes, adaptive capacities and reputation effects. The realist basis of the argument is strengthened with the insistence on unpacking, first the characteristics of the systems and services provided, second the many different governance possibilities and third the mechanisms determining actual governance structures. In short, the view of the organisation of IT provision taken here is an abstraction

that emphasises governance structure in terms of certain mechanisms informed by extant theory. This abstraction constitutes in itself a synthesis of various more partial and higher level abstractions, that can as a whole be re-synthesised with other factors (such as intra-firm politics and the bandwagon effect) that have been temporarily neglected in this thesis.

The normative presentation of the proposed framework in chapter 5 is also made on the basis of potentiality. There are no concrete 'cookbook' instructions. Instead, the various mechanisms included are potentially relevant factors in setting up the governance of IT provision in a particular situation. The assessment of the actual relevance of each mechanism and of its interplay with other mechanisms and contingencies are necessarily left to the decision maker. The empirical work presented in chapters 4 and 6 provides some insights as to the appropriate placement and scope of the postulated governance mechanisms within the wider context of intra-firm operations, inter-firm exchanges and IT provision and use. The following section turns on the empirical work carried out.

6. Empirical methods

As mentioned above (section 4), the aim of this thesis is to provide a theoretical account of the governance structure of the provision of IT services, which is based on the realist philosophy of science and which has a normative character with a view to aiding decision makers in practice. The realist and normative aims have been pursued with intensive case studies. One long-term case at British Petroleum, akin to participatory action research, has lead to the development of the proposed framework. Two shorter case studies (at ICI and Anglian Water) have been employed in retrospect to assess the proposed framework in alternative contexts, to examine its relevance and explanatory power and to reposition it within a broader web of governance determinants. The realist and action perspectives are considered in turn in the following sections.

6.1. Realism

Given the realist ontology summarised in Table 3.1, Sayer (1992) has proposed a classification of four types of research, which is reproduced in Figure 3.1. Abstract research deals with the structures and mechanisms of objects in the real domain with little or no formal empirical grounding. This would be the case with theoretical modelling in economics. Intensive research starts from a narrow slice of observations in the empirical domain and attempts an in-depth study in order to postulate real objects and generative mechanisms. Some case studies, action research, ethnographic or anthropological studies provide examples of methods in this category. Extensive research tends not to involve abstraction into the real domain and searches for regularities within large collections of empirical observations. Surveys and the corresponding data analysis methods are the typical examples of this type. Synthesis is a type of research that attempts to explain major parts of whole systems by combining extensive analysis of both empirical observations and of postulated abstract structures and mechanisms. Clearly, this type of research confronts practical difficulties and requires substantial resources.

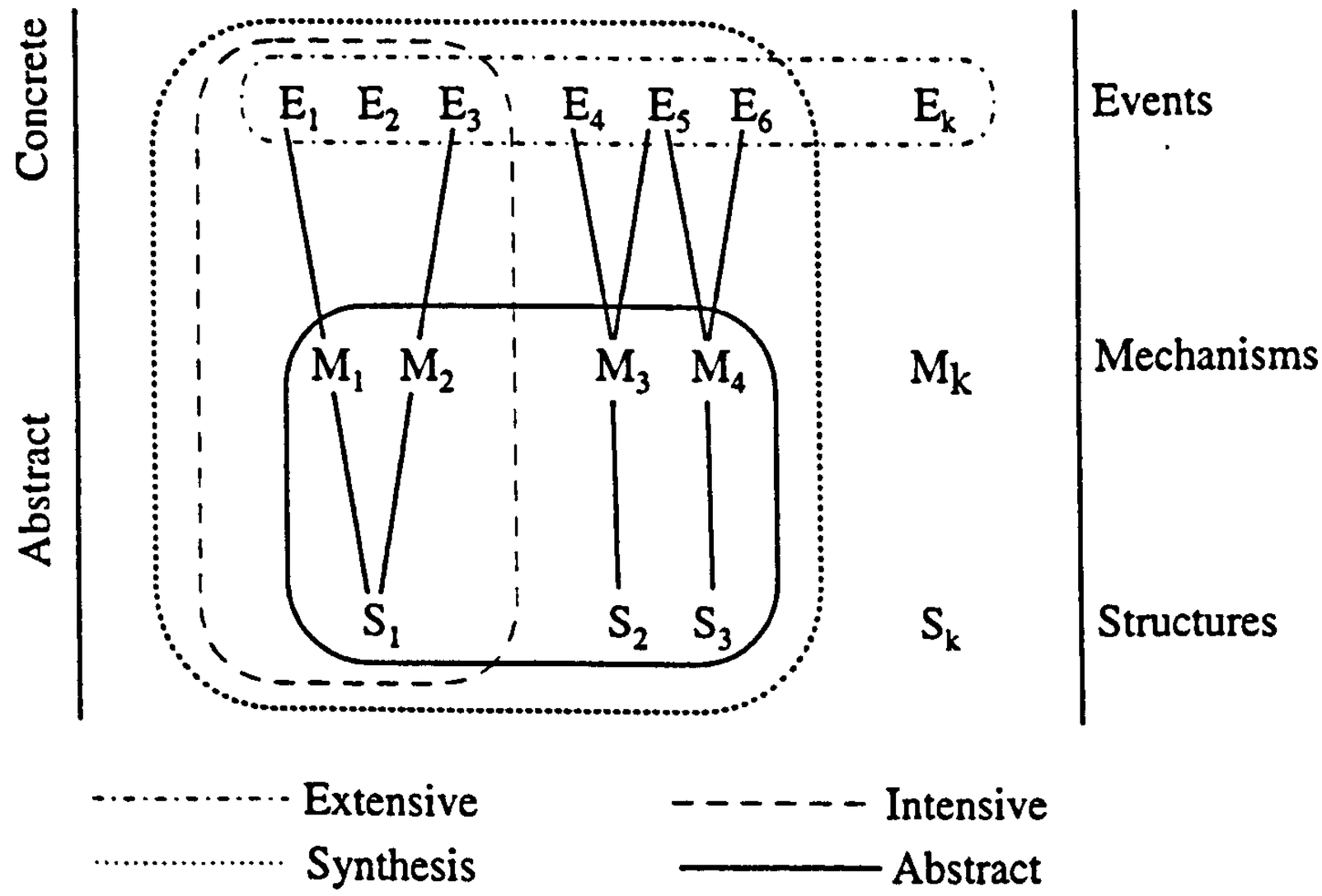


Figure 3.1

Types of Research from a Realist Philosophical Perspective
Amended from Sayer (1992)

Intensive and synthetic research are the most valid types from the perspective of the realist ontology and epistemology (Tsoukas 1989). Given that causality and necessity are not to be found in event regularities, the latter are neither necessary nor sufficient for theoretical explanation. Moreover, the ability to generalise can only be based on accounts of intransient real objects rather than on empirical coincidences. On the other hand, purely abstract research lacks the necessary empirical grounding which is the only source of validation.

The present research project has both an intensive and a synthetic dimension. Synthesis has been pursued at the purely abstract level by combining a broad range of theoretical propositions from the economics literature. Empirical grounding has been sought intensively through the study of IT outsourcing practices in the three companies.

The case study method is very versatile and can be employed in many different ways, including a multitude of data collection techniques (Eisenhardt 1989, Cavaye 1996). In information systems research it has been used extensively to serve positivist, interpretivist and critical research programmes (Myers 1997). The main advantage of case studies is that they allow researchers to study the people, the social structures and the phenomena they are interested in, in their natural setting (Benbasat et al 1987, Silverman 1998) and to explore qualitative nuances of the context and of the subjective understandings of the participants. Researchers are advised to enter a case without prior conceptions of the boundaries of their investigation or the constructs they need to study. These emerge and are constantly adjusted as the case study unfolds (Pettigrew 1985, Benbasat et al 1987).

As such, it is argued that case studies are best suited to the exploratory stages of a research process, serving to clarify the research problem and to inform the formulation of hypothesis, which subsequently feed typical extensive positivist testing. However, it has also been argued that case studies can be used for purposes of hypothesis testing (Yin 1989, Benbasat et al 1987), despite their weaknesses,

such as lack of experimental controls and small number of observations making generalisation problematic from a statistical point of view (Galliers 1991). In this positivist view of the method, single case designs are considered suitable for exploratory or disconfirmatory purposes or for exceptional (unique) cases, while multiple case design is considered suitable for theory testing (Yin 1989).

The realist philosophy rejects some of the positivist requirements for case-based research design. In particular, the number of sites is not considered to impinge on the generalisability of the theory (Tsoukas 1989) and the criteria for validity and reliability are reconsidered (Numagami 1998). Moreover, once we "*recognise the relative and multifaceted nature of truth amongst those in the research process*" (Pettigrew 1985, p. 57), it becomes clear that the objective formality of the positivist process of theory development and testing is not the only way to conduct case studies. However, this is not and should not be an excuse for lack of systematicity and rigour (Eisenhardt 1991).

Two different kinds of case studies have been carried out for this thesis. First, a single case study at BP aimed at investigating IT outsourcing practices through the lens of economic theories of the firm. In particular, the case served both exploratory as well as explanatory purposes and lead to the proposed framework. Second, two shorter case studies served the assessment of this framework in alternative contexts.

The diverse practices and opinions within the BP group have provided the stage for understanding the phenomenon of IT outsourcing beyond the secondary accounts of the literature and for deriving the proposed framework. Since all research is theory-laden, economic theories of contract and governance have been consciously adopted in this project, mainly due to their intuitive appeal in questions of 'make or buy' and to their prior use in IT outsourcing research (see chapter 1). In tandem with the principles of the realist epistemology (section 3 above), the purpose is not to confirm or falsify the theories but to investigate their relevance and to reframe

their applicability in an organised framework, accepting that they offer valid and useful, even if partial and incomplete, explanations of governance structures and contractual practices. The relevance and applicability of the various economic theories have been assessed over a period of nine months in an interactive, participative manner with managers at BP. In the process, practical experiences and opinions were exchanged with theoretical insights. The former were used to inform the selection and interpretation of relevant theoretical propositions. The latter were debated with BP management in an effort to understand their relative importance and applicability. The framework of chapter 5 emerged as a result. The practices and opinions observed at BP are documented in chapter 4.

After further refinement, the proposed framework guided data collection in two shorter case studies at ICI and Anglian Water. The aim there was to assess the adequacy of the framework in describing, understanding and explaining the observed practices in two contexts different to the original BP environment in which it was conceived. In doing so, some explanatory limits were also highlighted. Those parameters that were found to impinge on the observed practices but were not taken into account by the framework helped to identify its relative position within a broader web of causal powers in the organisational and economic context. The proposed framework and the theories it relies on are not considered as competing against alternative explanations such as political, evolutionary or resource dependence theories, as they are sometimes portrayed (Lacity and Hirschheim 1993b and Pfeffer 1997 [p.192]). Following a realist epistemological perspective, it can be argued that these theories are complementary, highlighting different aspects of the same phenomenon or explaining it at different levels of abstraction. The two cases and the respective results are the subject of chapter 6.

The utilisation of the three case studies is consistent with the realist ontology and epistemology. Of course, like every social theory the proposed framework is limited to the time-space window in which it emerged, namely large private companies in the UK today. As such it is a tentatively valid theory, subject to further refinement

and evaluation in alternative empirical contexts. A further note on the validity and reliability of the proposed framework is made in chapter 7.

6.2. Action

Within most management disciplines and in information systems in particular, the practical usefulness of theoretical contributions features prominently in debates among members of the academic community. Discussions on rigour versus relevance are one manifestation of such concerns. The editorial policy of MIS Quarterly gives almost exclusive priority to research articles that offer direct practical insight, alongside their theoretical and methodological rigour. In a similar vein, the research project presented here aims at delivering a theoretical account that can be of immediate use to managers deliberating the governance of IT provision in their organisations. For this reason the proposed framework is presented in normative terms. The case study at BP has been pivotal in incorporating such an element of action within the theoretical framework.

A notion of a theory aspiring to be both a valid explanatory account as well as an instrument for guiding decision making in practice, points to the method of action research. The term was coined by Kurt Lewin in the 1940s in the field of social psychology (Kemmis 1988). Over the last fifty years, it has been widely employed in community development, education and organisational psychology. It has also found fertile ground, but not widespread use, in operational research (e.g. Hindle et al 1995) and information systems, typically in research on systems development (Checkland 1991, Wood-Harper 1985, Sandberg 1985, Baskerville and Wood-Harper 1996, Lau 1997).

In its original conception (Rapoport 1970, Clark 1972, Kemmis 1988), action research aims at contributing both to the solution of practical problems of people in a particular problematic situation and to the advancement of social science equally, by joint collaboration between researchers and practitioners. Action research

entails a spiral process of planning an intervention, implementing it, monitoring it and reflecting on the whole process.

However, the practice of action research has not sustained a consistent definition, a concrete method or a particular epistemological foundation (Kemmis and McTaggart 1988). Today there appear to be several variations on the original theme without a well defined mainstream (Elden and Chisholm 1993). One fundamental strain pertains to the relative emphasis that should be placed on action (i.e. solving practical problems) or on research (i.e. general scientific knowledge). Confronted with these ambiguities, action science (Argyris et al 1985) distances itself from action research by making specific and strong claims on its definition, although the central aim is identified with Lewin's original ideas.

Given the ambiguity surrounding action research, it is too easy for any research project containing a practical perspective to label itself as such. To be sure, the project contained in this thesis has a strong element of action but one which is markedly different from both action science and the widely held notion of action research.

The case at BP was an action research project because it responded to an immediate need of management to improve their understanding of IT outsourcing, faced with the problems and uncertainty that ensued after the large deals they had arranged at the time. A framework to help the management at BP Oil understand and assess their IT outsourcing options was their expectation from the very beginning and throughout the process. The process was participative and interactive, with BP's experiences contrasted to theoretical insights until a common understanding was reached. BP management adopted a considerable amount of theoretical terminology, while their experiences and opinions were crucial in shaping the final

framework. In the end, the latter was adopted by BP and featured in their internal discussions and documentation[3].

As such, the project carried out at BP (which corresponds to the derivation of the proposed framework) is an action research project. The practical problem was to improve understanding and the project succeeded in doing so. However, the framework has not been tested in a full scale reorganisation of the provision of IT. BP had recently concluded their outsourcing agreements and no major changes were planned or intended. This is a potential source of limitations to the actionability of the framework[4]. For this reason and in order to avoid confusion given the ill-defined notion of action research, the term is avoided in favour of the more general notion of case study. The subsequent case studies at ICI and Anglian Water were not concerned with actionability, although the framework was presented to and briefly discussed with management.

An important difference between this project and a typical action research exercise concerns its epistemological grounds. In particular, action research and action science alike, take a pragmatist stance. As briefly mentioned in section 3 above, pragmatism entails the choice or development of a theory that suits the actual circumstances encountered, rather than conforming to any firm ontological claims. As Argyris et al (1985) successfully point out, this has an obvious advantage in that practising managers, being non-scientists and having limited capacity to perform lengthy analyses when rapid decisions and actions need to be taken, require brief and concise knowledge that is immediately applicable. Such knowledge, according to the deductivist argument of Argyris et al (1985), cannot be but of the contingent

---

[3] For example, OHP transparencies used at internal presentations in BP used the language and displayed various aspects of the S-CAGE framework. In one transparency asking why did BP's outsourcing arrangements not go well in certain respects, the first bullet-point answered "because we didn't follow Nikos' framework".

[4] This does not constitute a limitation to its explanatory power. The latter is discussed in the previous section.

form "if action x is taken, consequences y will result". Pragmatism serves this purpose by quickly reducing a potentially complex theory to few such statements according to what is perceived as relevant and/or important given the circumstances and the goals of the practitioners.

Indeed, management at BP regularly asked for such concrete guidelines, in particular of the form "service x is best procured under governance structure y". As it has been shown, however, such claims are illegitimate both for philosophical (see section 2 above) and for substantive reasons (see last section of chapter 2). BP's requests were resisted. The shape of governance is determined on the basis of the contingent (and hence unpredictable) conditions arising in each individual case.

However, the argument of Argyris et al (1985) is still valid and the question remains: how can managers with limited information processing capacity accommodate a potentially complex theory in an 'on-line' manner? As the same authors suggest, the managers ought first to be trained to understand and apply the theory. The way the framework of chapter 5 is presented as an organised collection of potential governance mechanisms facilitates the reduction of the many elements it contains to those that are relevant to a given actual context, by a knowledgeable manager. This became obvious in interviewing managers for the case studies at ICI and Anglian Water. The open ended questionnaire employed (see appendix) served as a guide through the elements of the framework. During the interviewing processes it became clear that the number and complexity of the issues raised in the questionnaire quickly reduce to fewer and more specific questions that become relevant once the facts start unfolding, while many alternative possibilities were quickly eliminated as less relevant or less important.

Therefore, the proposed framework maintains its practical usefulness (actionability) by virtue of the empirical exercise at BP and of the way it can be reduced to fewer relevant recommendations, while not compromising its adherence to the epistemological requirements of the realist philosophy of science.

7. Conclusions

The research design explained in this chapter is an effort to synthesise and respond to the epistemological and methodological concerns raised in the information systems literature (e.g. Fitzgerald et al 1985). The limitations of the typical quantitative study that follows the principles of logical positivism is increasingly being recognised (e.g. Klein and Lyytinen 1985, Pervan and Klass 1992). A greater plurality in the choice of both philosophical assumptions and particular methods is also being advocated (e.g. Galliers 1992, Myers 1997). Among the research methods, the case study, in its various mutations, has become quite popular alongside more 'mainstream' quantitative methods (e.g. Cavaye 1996). Among the philosophical assumptions, interpretivism and, to a lesser extent, critical theory are gaining ground (e.g. Orlikowski and Baroudi 1991, Walsham 1995).

This thesis takes on board the above issues and responds both in terms of philosophy and in terms of methodology. First, philosophically this thesis makes a modest contribution by introducing transcendental realism and outlining the latter's implications for theory development and empirical methods. These principles govern both the research methods as well as the substantive theoretical propositions. Second, methodologically the proposed framework is developed through an intensive case study with strong elements of participatory action research. Two additional case studies are employed to assess the proposed framework in alternative contexts. On substantive matters, in order to avoid re-inventing the wheel, economic theories of governance have been employed as a basis for theoretical development and for the proposed framework.

By making a relatively brief reference to philosophical issues in social science, as they are viewed through the philosophical lens of transcendental realism, this chapter attempts to clarify the nature of the project undertaken, its ends and its means. It may appear as a rhetoric for the research approaches followed in this

thesis and against other alternatives. It *is* a rhetoric inasmuch as it makes explicit the foundations and assumptions guiding this work, instead of leaving them to the reader to infer (Metcalfe 1996). In addition, however, it plays an instrumental role in three respects.

First, this chapter explains the way in which the economic theory of chapter 2 is subsequently used in the proposed framework. Given the realist epistemology adopted here and the positivist tradition in economics, the economic theories employed have to be reconceptualised as potential structures and mechanisms which may or may not produce observable effects of various kinds, depending on actual circumstances. Second, by addressing all these meta-theoretical issues, this chapter sets out the basis on which to judge both the substantive theoretical contribution of the thesis and its methodological sufficiency.

Finally, this chapter is pivotal in the presentation of the thesis. Chapter 1 set the motivation for undertaking research on the governance of IT and highlighted the main strands of the received knowledge from the extant academic and trade literature. The conclusion sketched the direction followed in the rest of the thesis for studying the outsourcing phenomenon and the organisation of the provision of IT systems and services. Chapter 2 set out to review the relevant parts of the economics literature and to elaborate a starting point for further theory development. The present chapter establishes the meta-theoretical position and the empirical methods for the substantive work that follows in chapters 4, 5 and 6, starting with the case study at BP.

CHAPTER FOUR

The Problem Facing Practitioners: A Case Study In British Petroleum Plc.

1. Introduction

The economic theories of organisation reviewed in chapter 2 are clearly pertinent to the governance of information technology service provision but the exact way in which they apply and their relative importance are not straightforward. The in-depth case study carried out at BP aimed at facilitating the transition from the theory of economic organisation and apply it to the question of sourcing IT products and services.

This case study serves to generate awareness of the practical issues specific to the provision of IT. More than that, the case study served as a 'distillery' of the theories in chapter 2. The end product is the S-CAGE framework presented in the next chapter. As explained in chapter 3, section 6, the case study had a strong action orientation. BP expected specific practical results from the exercise and contributed actively to the development and refinement of the framework. In particular, theories, meanings and practices were debated, and potentially useful frameworks were developed interactively with management at BP. For presentational purposes, the practices observed at BP are separated from the framework that was produced in the end, although the observation and the development of the framework progressed simultaneously. This chapter presents facts and opinions collected at BP. Chapter 5, in turn, combines the relevant theoretical insights with the practical experiences to present the new framework.

This chapter documents events, contractual arrangements, processes and opinions with a view to understanding how the mechanisms of governance are manifested in practice and what consequences they have. The following section describes the intensive interaction process carried out at BP, its strengths and weaknesses. Section 3 sets the scene at BP as it relates to organisation and information technology. Section 4 discusses the diversity of opinion across the BP Group with respect to outsourcing. The supplier selection and the transition process followed by BP Exploration are the subject of sections 5 and 6 respectively. The structure of different outsourcing contracts is considered in more detail in section 7. Finally, some further issues of concern faced by BP managers are presented in section 6 and the chapter is concluded in section 9.

2. Carrying out the case study

The case study was carried out at British Petroleum Plc from May 1994 until January 1995. The process involved collection of written documentation as well as interviews, presentations and discussions with a number of IT managers across the organisation. The Global Head of IT of BP Oil at the London headquarters and his associates were the main points of interaction.

Throughout the process, the goal of the meetings was twofold. First, to collect qualitative and quantitative information consisting of factual (oral and documented) data and the managers' perceptions about IT sourcing problems and practices in the different BP businesses. Second, to inform and discuss with BP managers concepts and ideas derived from theory and debate their relevance and applicability. In this way, the case study became a learning process for both the researcher and the participating managers. The former familiarised himself with the language and frame of mind of the practitioners and the latter became able to access and adopt complex theoretical concepts and jargon. In the beginning there was a clear communication gap, due, most likely, to the different language used, that of the academic and that of the practitioner. However, over time the parties

converged to common grounds of understanding as both sides kept bringing into the discussion new information, ideas, concepts, as well as challenging feedback. This does not mean that a consensus of opinion was reached. On the contrary, participants maintained quite diverse and strong views. However, it took a while to bridge the gap between the researcher's academic language and terminology and the practitioners' own use of outsourcing terminology, which had been developed through their specific experience in BP.

Most interactions with management at BP started as interviews but evolved into collaborative exchanges of ideas. As mentioned below, interview guides (open-ended questionnaires) were prepared for each occasion. The chosen theoretical background (chapter 2) informed all interview guides, discussion papers and debates. In particular, many interview questions were directly related to specific theoretical propositions. In other instances, theoretical propositions were explained in simple terms to practitioners and were then debated in order to assess the extend of their practical relevance. In doing so, two methodological aims were pursued. First, this approach provided a link between the *abstract* mechanisms postulated by theory and their *actual* manifestations. This relates to the stratified view of reality, advocated by the realist philosophy of science (see chapter 3, section 6.1). Second, this process of feeding theoretical concepts to practitioners and debating them educated the latter who, in turn, were able to reflect further on their experience. This served the action research objectives of the case study (see chapter 3, section 6.2). More specifically, the interaction with participant unfolded as follows. The information collected and the ideas generated at each meeting provided the basis for revising and targeting the interview guides for subsequent meetings. Meticulous notes were taken during meetings and many issues were revisited with the same or other participants in other meetings. After each meeting, all information (notes and documents) was analysed, interpreted and drafted in the form of interim discussion papers, which were shared and debated with participants in the next meeting. In this way, the whole process was interactive and augmented the knowledge and understanding of both the managers and the researcher. The former reflected on

their practices and revised their beliefs through a new perspective. The latter was able to pinpoint the practical relevance and manifestation of abstract theoretical propositions. The aim was to generate a unifying framework that would provide a common ground for understanding and bridging the diverse opinions and preferences held by each participant. Clearly, the final framework does not explain absolutely everything. Its limitations are discussed in chapters 5, 6 and 7. Nonetheless, it is the product of an intensive, evolutionary and collaborative process of consolidating theoretical propositions, subjective opinions and organisational and contractual practices.

During the process, the researcher had the initiative for analysing the data, proposing alternative frameworks and seeking criticism and feedback. The hosts at BP Oil demonstrated commitment in drawing support for the study from all quarters of the organisation. Below, the process of carrying out the case study is described in chronological order. Thus, it becomes evident how the aims of actionability and realism (chapter 3, section 6) were achieved.

May 1994

The case study started with a visit to BP's corporate headquarters in the City of London. The main event of the day was a presentation to seven senior and middle IT managers from Oil (the Global Head of IT and two advisors), Exploration (two senior IT managers) and the Corporate Centre (two desktop services managers), followed by person-to-person fact finding discussions[1]. The presentation was based on a paper titled 'the impact of information technology on buyer-supplier relationships', distributed to the participants. The paper outlined the main tenets of transaction cost economics and agency theory and discussed their applicability to IT outsourcing. To exemplify the points being made, the three-dimensional contingency framework proposed by Mahoney (1992) formed the basis of

discussion. The terms 'governance', 'specificity', 'programmability' and 'separability' attracted much interest, despite their abstract, academic, nature. The ensuing debate revolved around understanding the framework and associating it with the experience of BP. Moreover, the potential direction and outcome of the research and a clarification of the expectations of either party were openly agreed.

In particular, and in accordance with tentative prior communication, it was agreed that BP would contribute information, documentation, management time for interviews and feedback on the development of the anticipated framework. In return, they expected to gain a better understanding of the issues involved in IT outsourcing and of the range of possible arrangements[2]. It was agreed that the final deliverable would be a framework classifying various governance alternatives and the conditions under which they are applicable. The framework would have to have practical usefulness for BPO but its exact form was, at that stage, left open.

Two points in that agreement highlight the action research aspect of the case study. First, BP was expecting a tangible benefit in the form of 'education' in new ideas and in the form of a managerial framework. Second, BP would contribute not only information but also active feedback on the development of the framework. In this sense the case study was to be a catalyst for transforming abstract economic theories into actionable knowledge. An electronic mail message that was later

---

[1] As a result of recent reorganisation throughout the BP Group, formal managerial titles were being abandoned in favour of ad-hoc specification of responsibilities (see also Cross et al 1997). This is why generic terms such as 'advisor' and 'senior manager' are used.

[2] The research was not going to address all the issues involved in IT outsourcing but only those afforded by the transaction cost and agency theoretic perspective that had already been adopted by the researcher. BP was also receiving other academic insight from within the UK and from abroad. BP felt that the research proposal considered at that meeting was making a distinct contribution. The recent outsourcing agreements of BP Exploration had spawned a debate within the Group on the appropriateness of various sourcing arrangements (see also section 4 below). Therefore, there was a desire to reconcile the different opinions under a unifying framework.

circulated to enlist managers for meetings and interviews explained: "*[...] He aims to identify appropriate governance structures [...] for various IT services [...]. For our part, we are providing Nikos with our experience of outsourcing and a sounding board for his ideas.*"

June 1994

Meetings were held with the two Oil IT advisors during a single day, both individually and jointly. The purpose of the meetings was mainly to collect facts about the history and structure of IT sourcing practices in the Group. In addition, various issues were debated.

A discussion paper, prepared before the meeting, was shared with the two managers. It included an open-ended interview guide as well as brief summaries of issues for debate. The questions asked were related to the various sourcing practices in the Group and the recent BPX contracts. Issues for debate included the irreversibility of outsourcing decisions, the specificity and complexity of systems, the distinction between core strategic versus tactical activities, the choice between one (or few) versus many suppliers and the notion of partnership vis-à-vis inflexible contracts.

The purpose of the issue debates was to elicit the opinions of management, the actual practices of the company and their relative importance and relevance in the context of BP. The questions and issues in the paper were drawn from economic theories, from the IT outsourcing literature (which, at the time, was limited but growing) and from the discussions during the previous meeting.

After the meeting, a report was prepared and sent back to them for circulation to interested parties in BP. The title of the report was 'Information technology outsourcing: an economic perspective on the nature of services and efficient contracting' and its aim was to educate and to raise awareness within BP of the perspective and of the terminology. It was expected that BP managers would be

able to associate the theory to their experience by themselves and reflect on its relevance during subsequent interviews and meetings. Once more, the paper, based on Mahoney's (1992) framework that was presented in the first visit, explained the terms and their potential applicability by giving relevant examples. Although the report did not contain immediate feedback on the information thus far collected, it gave stylised examples, drawn from the BP experience, to clarify the terms.

August 1994

After a break, the case study resumed with a meeting with one of the two Oil IT advisors. More facts and documentation were collected and plans were made for further interviews. The possible shape of the framework was also considered. These ideas, alongside other empirical findings and theory-driven arguments were, after the meeting, elaborated on a series of slides and a new 16-page document with the same title as the one before. Its purpose was to communicate and debate these ideas in subsequent meetings. This paper, which since then was continuously being revised, was the first draft of the final report[3], although very few of the original ideas survived to the end.

It is worth highlighting some of the elements of that paper. It introduced the notion of 'service fulfilment devices'. These were defined as contractual and organisational arrangements aimed at ensuring the achievement of the desired goals in terms of service delivery and comprised three categories: measures of performance, monitors of conduct and formal incentives. They were associated to the three variables of Mahoney's (1992) framework, namely task separability, task programmability and asset specificity respectively. The concept of a lifecycle of service fulfilment devices, which survived nearly intact to the present thesis, was also introduced.

---

[3] The final report presented the S-CAGE framework and formed the basis for chapter 5.

The paper proposed a hierarchical classification of IT services, discussed possible forms of service fulfilment devices for IT services at different levels of the hierarchy and explored the cost (difficulty) of developing them through the lifecycle. Finally, the paper considered the advantages and disadvantages of organising IT services under a single unified governance or through a portfolio of governance structures. All these ideas were the seeds of the elements of the framework presented in chapter 5.

However, at that time a different argument started taking shape. The argument, which was also appealing to BP, was that a particular set of service fulfilment devices could be devised for each IT service. The set of service fulfilment devices would then dictate acceptable general forms of governance. This blueprint for a contingency framework was contemplated for about a month before it was abandoned. It was considered to be too complex and too ambitious. It was also thought that it was not really necessary to be so prescriptive.

Later in August, this paper was discussed at length during further meetings with the two IS advisors and the Vice President for Operations of BP America who had visited London to present their approach to IT sourcing. Several good ideas were exchanged that influenced the final framework, including, among others, the importance of proper definition of services and the varying degree of management involvement in monitoring the contract (and the provider) that is required for different services.

September 1994

Whereas all previous meetings were held at the corporate headquarters in London, on 1st September various meetings took place at the Stockley Park headquarters of BP Exploration and at the Hemel Hampstead Headquarters of BP Oil. A revised interview guide was prepared for these meetings. Related documentation was collected from both sites, including internal policy documents, vendor performance reports and parts of contracts. A senior manager responsible for Exploration's

global IT infrastructure explained in detail the initiative, the rationale and the structure of their recent contracts. At Hemel, the Partner Resource Manager (PRM - see section 7.2 below) for Oil Europe explained the 'grief point' system implemented in the respective contract with SEMA. The part-time PRM for Air BP at the same site gave similar information regarding Air's contract. On the same building, SEMA rents space for their local operations (desktop support, network management and related services). SEMA's Account Consultant (local site manager) provided the vendor's perspective. Among others, he stressed the importance of additional (out of scope) work for SEMA and he explained their perception of value-added contribution to BP.

Apart from giving factual information, all commentators that day were asked to critique the implementation of the contract and the effectiveness of the performance measurement systems. Several interesting points emerged regarding, among others, the form of supplier and customer opportunism, the qualities (precision and completeness or the lack of them[4]) of performance measures and the meaning of 'partnership'.

Later in September, a meeting was held with SEMA's Account Manager responsible for BP, at the SEMA headquarters outside Birmingham. A separate interview guide was prepared for this occasion and further contract documents were collected from SEMA. Among other issues, the discussion emphasised the contract management procedures, the degree of customisation of contracts to different clients and services and the partnership relation with BP.

November 1994 onward

Meanwhile, the paper mentioned above was undergoing major revisions. By the end of September, it was still under the same title but it was substantially enhanced and extended. The contingency approach contemplated in the previous version

---

[4] See chapter 5, section 4.1.2.

had been abandoned (see chapter 3, sections 4 and 6.2 for a more extensive discussion), in favour of a classification of the elements of governance. The three-layer model (chapter 5, section 4) had started taking shape and formed the largest part of the report. In addition, the paper contemplated various arguments regarding 'total' outsourcing solutions versus narrowly focused contracts. On 7 November, one of the two BP Oil IT advisers visited Warwick Business School for a review of the drafts and an update on relevant developments at BP.

By the end of the year, the governance continuum (chapter 2, section 8.1) and the idea of service clustering (chapter 5, section 3) were introduced. Among other explanatory notes, the final report included a generic methodology for applying the framework in practice. On 15 January 1995 it was presented to and discussed with a panel of IT managers from Oil and Exploration at the London headquarters. The new title of the paper was 'The efficient governance of information technology service provision: The S-CAGE framework'. Most of the managers who participated in the interviews were present. All participants received the framework with great interest. Feedback from the presentation lead to a revised version of the report and a supplement titled 'A guide to understanding IT service specificity'. Both were distributed widely within BP in early February (see chapter 3, section 6.2 for some indication of the positive way in which the framework was received in BP).

In May 1995, Oil's Global Head of IT and a senior member of his team at the corporate centre[5] visited Warwick Business School to discuss further developments on this project. While the case study was effectively completed, close contact was maintained. As the research and the writing of chapters 4 and 5 progressed, the report was continuously being updated and refined using further feedback from BP Oil. Soon, the title was changed to 'To outsource or not to outsource? This is *not* the

---

[5] By that time the Oil Systems Group at the London headquarters had been restructured more than twice and had undertaken responsibility for the systems of the corporate centre. Of the four advisors working at that level in BPO and BPCC when the case study had started a year before, only one had remained.

question: How to organise the provision of information technology'. The last version was produced in January 1997.

2.1. Strengths and limitations of the approach

The case study was a transformation process. Inputs were the recent BP experience with IT outsourcing and the range of economic theories reviewed in chapter 2. The output is the framework presented in chapter 5. BP had entered into large outsourcing agreements for a range of different services. This change had a significant impact on the IT organisation within all businesses of the Group. BP contributed to the case study with diverse practices, opinions and reactions to the outsourcing exercise. The theories selected are also characterised by extensive diversity. The literature surveyed in chapter 2 spans a broad range of theoretical propositions. The transformation process that took place during the case study was synthetic in nature. It synthesised the diversity of experiences at BP with the diversity of the respective literature to produce a single unifying framework. The BP experience served as a catalyst for the relevance, applicability and practical manifestation of the abstract theories. In turn, the theories provided the means for reconciling the diverse empirical observations onto a common basis for understanding them.

Like every empirical research, this case study has its own strengths and limitations. One strength of the case study is its emphasis on practical relevance while maintaining strong theoretical grounding in the literature. The researcher and BP traded theoretical insight with practical experience in the process of converging to a common ground of understanding (the final framework). The collaborative interaction with practitioners was a key feature of this process. It has both advantages and disadvantages. The main advantage is that it facilitated the aims of relevance and actionability. The main disadvantage is that it was susceptible to biases emanating from the researcher's theoretical predisposition, from the specific experience of BP and from the personal and political inclinations of the

participants. In particular, the process emphasised the governance aspects of outsourcing at BP and has thus set aside other, possibly important, organisational antecedents of the outsourcing decisions and the ensuing events. Furthermore, the generality of the findings is compromised by the fact that the source of evidence is only one organisation. Additionally, some of the in-house outsourcing terminology used by BP managers has been adopted in parts of the S-CAGE framework (chapter 5). This is not a serious problem, as the essence and validity of the framework is not affected. For example, chapter 6 uses the terminology of the companies referred to there, without contradicting or causing any other misinterpretation of chapter 5. Having said all that, BP demonstrated significant diversity of opinion (among individuals) and practices (between different companies in the group). However, some of this diversity might be due to conflicting interests and other political interactions among individuals in BP. This possibility (see, for example, Lacity and Hirschheim, 1993) has been left out of the scope of this research.

As discussed in chapter 3, any research project is theory laden and both researchers and practitioners bring their own subjectivity to the research process. Moreover, the research process and the research outcome themselves feed back into social reality and reconstruct it by influencing the beliefs and actions of social actors. Research approaches adhering to the principles of experimentation emphasise the need to control all such biases in order to obtain objective observations of cause and effect. Interpretivist approaches insist that not only are subjective beliefs part of social reality (and, therefore, they should be part of the study) but also that such biases ultimately cannot be avoided[6]. In this case study, no effort was made to disable these subjective interjections. On the contrary, both the researcher and BP managers agreed openly from the outset to embrace each other's knowledge and use it to revise their own understanding and to construct a set of common beliefs.

---

[6] The various strands of interpretivism are not unanimous on this. See chapter 3, section 2.

Therefore, the result is likely to be biased by the specific experience in BP and by the chosen theoretical background. This possibility, among other things, is examined in chapter 6. As it is also mentioned in other parts of this chapter, it should be stressed that BP did not present a unitary case. A diversity of organisational and contractual practices was observed in different BP businesses. Moreover, opinions varied greatly between individuals. So from a purely experimental viewpoint, there has been some degree of randomisation. However, richness of context was a more important objective than randomisation.

The case study has lead to the refinement and extension of extant theory, in the form of the governance continuum (chapter 2, section 8.1), the notion of clustering and the classification of the elements of governance (chapter 5, Figures 5.2 and 5.4). To summarise, this case study has not been a test of theory and it has not led to the development of major new theory. Notwithstanding, it has been a research intervention in BP and it has achieved a synthesis of two diverse realms of knowledge, namely the abstract economic theories and BP's experience.

3. Organisational background

British Petroleum Plc. is one of the largest petroleum and petrochemicals companies in the world. It is organised in three main divisions under the control of the Corporate Centre (BPCC) and operates in eleven regions all over the world. Exploration (BPX) explores and produces crude oil and natural gas. Oil (BPO) refines, markets, supplies and transports oil products. Chemicals (BPC) manufactures and markets petrochemicals. Oil is the largest division generating nearly 75% of the Group's turnover which was over £36 billion in 1995. When the case study began in 1994, of those three businesses, Oil was the largest IT user with 70% of the expenditure and 80% of the headcount. Total IT expenditure of Oil for 1994 was $350 million. This represented a 22% reduction since 1989.

In 1994, BPO's model of IT procurement included three basic categories. First, in-house activities included service and project management, procurement, and policy. Second, equipment, software and commodity services were purchased. Third, data centres, telex/telephone and 24-hours operations were outsourced. The procurement of other services varied among business units and sites according to local conditions. For example, telecommunications and bandwidth were either purchased from a telecommunications operator and managed by BPO or wholly outsourced. Help desks, applications development and support and consultancy were either carried out in-house or outsourced. Total internal activities totalled about 39% of expenditure, purchased services accounted for 26%, facilities management for less than 2%, work contracted out represented 23%, fixed price contracts represented 8% and cost-plus contracts less than 2%. The last four categories are broadly considered as 'outsourcing' by BPO and constituted a total of 35% of IT expenditure. The bulk of infrastructure activity was either outsourced (in the broader sense) or purchased. Business applications were delivered by a roughly equal balance between outsourcing, purchasing and in-house work.

In BP Oil's terminology, 'purchasing' refers to equipment as well as to fixed-term contract staff. Facilities management is linked to the traditional notion of data centre operations. Contracting out refers to individual project-based contracts. Their notion of true outsourcing was of a bundle of services contracted out for a fixed price schedule, for a given period of time. At the time of the case study, BPX had already entered a long term total outsourcing agreement of the 'cost plus' kind. Some of Oil's IT (less than 2%) was included in this contract.

Oil's adoption of outsourcing coincided with a technical platform shift from mainframes to client-server. Linked to this was a major IT development which began in 1990 (before the outsourcing deals), the European Systems Project (ESP). This involved the development of integrated information systems for all functional areas covering the whole of Europe for Oil's operations. It is a very large project being developed at BP sites in the UK, Germany and France. The project was left out

of the outsourcing contract. Its development was the responsibility of internal groups. Contractors were employed for parts of the system in fixed and short term contracts. The organisation of ESP development is readily understood in the context of transaction cost economics and of those theories of the firm emphasising the organisation of knowledge (chapter 2). ESP required BP-specific knowledge and, given its innovative nature, it involved significant uncertainties which BPO could not forecast. It was also considered a project capable of bringing significant business advantages as it was designed to integrate all European operations. For these reasons BPO wanted to have close control over ESP. Since the case study, the project became international and was being deployed to sites throughout the world. Moreover, once the relevant business knowledge was embodied in the system, BP sought to reduce operating costs by contemplating more hands-off (discrete market) governance arrangements in collaboration with IT suppliers. Whereas specific knowledge and skills dictated a set of highly relational organisational structures during development (predominantly internal organisation), the structured and impersonal character of this knowledge after it is built into the system reduced the costs of organising it and thus gave the opportunity for the adoption of more discrete governance structures.

In 1991 BPX initiated a major outsourcing exercise which was concluded with a total outsourcing agreement with three suppliers, namely SEMA, Syncordia and SAIC. With each supplier BPX signed a framework agreement where the general terms of the relationship were detailed. Rolling call-off contracts were subsequently signed by each site or business unit. This exercise was initiated by and for BPX globally but in practice call-off contracts were activated mainly in the UK where they also included the corporate centre and over time expanded to include services for other businesses within the group.

The outsourcing exercise was part of the 'Project 1990' initiated in May 1990 by CEO Robert Horton which aimed at introducing major cultural, structural and managerial changes throughout the organisation in order to focus on BPs core

business and revitalise its financial situation. In the process, large numbers of staff were made redundant. IT was no exception and the outsourcing deal contributed in this direction. Within four years, the 900-strong group that provided all IT services internally to the corporate centre was reduced to just 7 in four main stages. The roles of the remaining staff are business consultancy, technology consultancy and contract management.

BP have been trying to standardise their systems and services across the Group whenever this is desirable and maintain non-standard systems only where they create substantial local value added. For example, they have a global and standard contract with a major software vendor for all their office productivity tools. Similar arrangements are in place for WAN services except for those countries where the telecommunications regulation is not convenient or where BP's main supplier does not reach. Most business applications, in contrast, are non-standard and different approaches to their support and maintenance may be followed by different business units. This practice reflects the trade-off between the business value added by asset specificity and the respective increase in management and transaction costs.

4. IT procurement philosophies in the BP Group

The BP Group is a heterogeneous organisation. Each business unit operates in different markets and follows a different business mentality. Their approaches towards IT procurement also differ. Nevertheless, while they are aware of their distinct characters, they also communicate and share experiences and best practices. Early in 1994, global representatives from the main BP businesses met to discuss their approaches to IT outsourcing, to identify the differences, to draw lessons for the future and to find some common ground. This meeting was seen more as a learning and sharing exercise rather than an effort to establish one uniform approach throughout the Group. Each business unit continues to follow its own path, according to its idiosynchracy.

The approach of BPX in undertaking the IT outsourcing initiative stemmed from the strategic imperative of the company to devolve functions, focus on core operations and slim down the organisation. BPO took a different approach. Their starting point was the internal efficiency of IT operations and this has been their guiding criterion in deciding how to organise IT procurement. The BPX initiative, however, acquired a critical mass that put pressure on other parts of the group to follow suit and to join the same framework agreements. Chemicals and Air BP (a subsidiary of Oil) joined soon after the first contracts of Exploration and the Corporate Centre.

This difference in approach reflects other differences between the two businesses. BPX is characterised by a more aggressive and proactive management attitude while in BPO changes are more subtly introduced. BPX is a highly centralised and hierarchical organisation as opposed to Oil which has a flat structure akin to notions of network organisation.

Oil has 2 senior IT managers in its global headquarters. Their primary job is to facilitate communication and diffusion of best practice across a global network of regional IT managers. Each region may follow different technical and managerial (including outsourcing) practices. A Global Professional Team (GPT) facilitates communication between regional management teams. The GPT only decides upon a few global standards, the adoption of which can generate cost savings or other advantages as compared to independent decisions and technology choices.

Perhaps more importantly, the systems of Oil and Exploration are very different. BPX has high technology systems and applications which are relatively independent and easier to define. Oil's business is much more complex, ranging from refinery and warehousing to distribution and retailing. The information systems supporting all these activities are both diverse and inter-related. It is not only more difficult to write comprehensive contracts, but the implications of any IT procurement decision for a given application are more far reaching than the case of BPX where IT

services are more distinct. A different approach to evaluating the outsourcing decision was therefore warranted.

BP America's philosophy for IT procurement is that as long as IT cannot be treated as a commodity service like electricity, emphasis should be placed on the definition of both the services and the set of skills required for managing their provision. Services differ in their characteristics so their definitions should also differ. In general, according to BP America's Vice President of Operations, IT service definitions should emphasise the business aspect of the service rather than the technical processes and assets involved. Moreover, the organisation should define the set of skills required to manage the delivery of services to customers, whether they are 'produced' in-house, contracted out or outsourced. Commodity services require from the organisation relatively narrowly defined skills for contract management while services unique to the business need a broader range of technical, business and legal skills. For example, a few people are needed to look after a telecommunications contract and the supplier's performance. With applications however, staff from the client organisation are needed to oversee the developer's effort and product more closely. The opinion of this senior business manager reflects an understanding of the repercussions of asset specificity and of the issues pertaining to input an output measurement. Of course, the terminology adopted in practice differs from that used in the theoretical arguments reviewed in chapter 2.

One of the underlying principles that guided BPX towards outsourcing was their decision that the IT group within the company should focus on business rather than technical solutions. They consequently wanted to separate information management from technology management. The practicalities of contracting (described later in the chapter) made them realise that the ability to separate business service from technology management varies among different services.

According to a senior manager, BP Exploration had three further reasons for pursuing outsourcing. First, they ascertained that the existing IT organisation could not improve its performance any further, at a time when the challenge within the company for continuous improvement was imperative. The outsourcing option promised to break through this performance barrier. From a fat hierarchy of 1400 people including many technical staff, they achieved a transition to a slimmer organisation of 160 people, of which very few (about 10) are technical consultants and most are IT professionals working with the business on their use of IT. The new structure resembles an inverted pyramid as shown Figure 4.1.

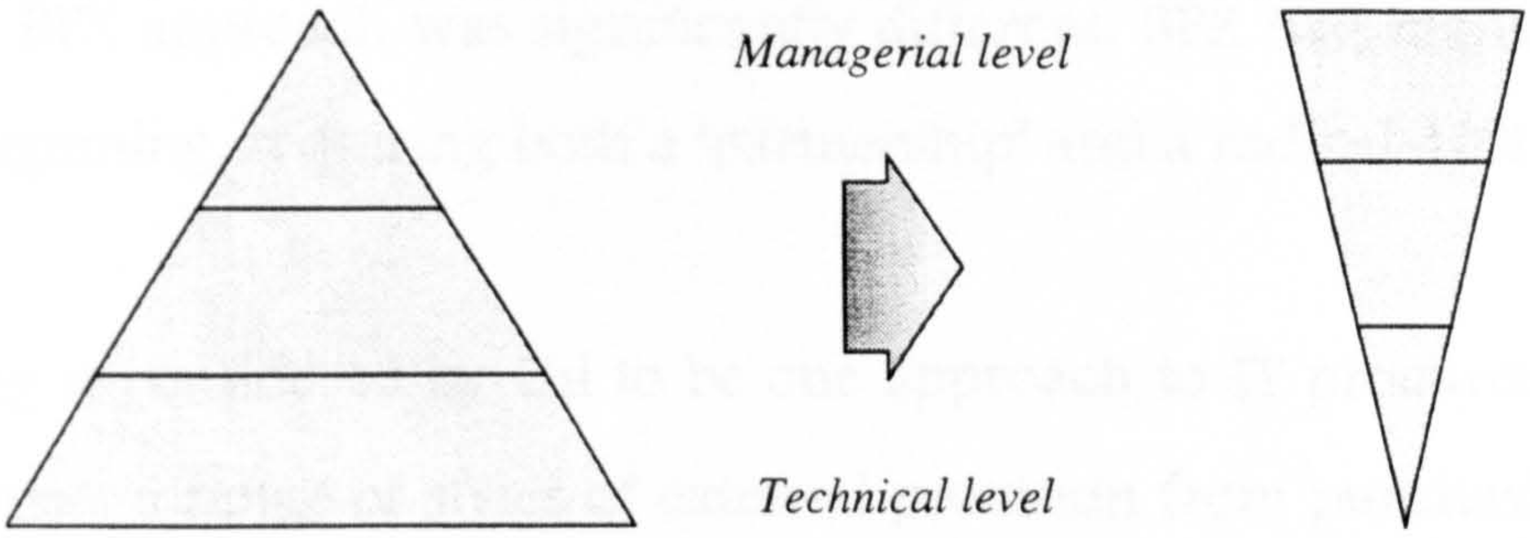


Figure 4.1

BP Exploration's organisational transition to outsourcing

From an internal hierarchy of 1400 people with fewer managers than technologists to a slimmer structure of 160 people with more managers than technologists.

The second reason for considering outsourcing was the desire to align the cost structure of IT with that of the rest of the business. BP Exploration's business is characterised by a large proportion of variable costs. Cost flexibility is consequently very important and outsourcing was seen as a way to increase the variable proportion of IT costs.

Third, the general business environment put pressure on BPX to reduce their total costs and outsourcing was seen as a cost cutting instrument. Another reason quoted by BPX was the problem of effort duplication and the difficulty of coordinating the large number of suppliers under the previous regime. After outsourcing they

interact with three major suppliers which may in turn use a number of subcontractors in a transparent way for BPX.

The BPO approach was incremental. The process of outsourcing IT systems and services at the Hemel Hampstead headquarters (outside London) started with a competitive bid to assist them with the transition from mainframe to client server. It was a 'closed book' competitive process and the contract was awarded to SEMA. According to the SEMA account manager, BP had relatively few expectations. This relationship however gradually evolved to a broader 'partnership' (in his own words) by including more services in the contract and implementing an open-book policy. The BPX approach was significantly different. BPX had higher expectations from the beginning, requiring both a 'partnership' and a radical 40% cost cut.

Outsourcing is considered by Oil to be one approach to IT procurement which in itself comprises a range of styles of external provision from purchasing through to contracting out using either fixed-price or risk-sharing contracts. Having realised there is a range of options for organising the provision of IT services, they have no *a priori* bias towards total outsourcing or towards a single supplier, despite the bandwagon effect created with the Exploration's initiative. The latter's attitude derives from their top level vision for IT which places emphasis on the 'demand side', i.e. the business value of IT, rather than the 'supply side', i.e. the process of delivering systems and services. For BPX there is virtually no alternative other than outsourcing. In addition, Oil seems to have a preference for fixed-price contracts while Exploration for cost-plus/risk-sharing contracts. Oil tend to select their procurement mode on a service-by-service basis and opt for outsourcing when it can deliver benefits beyond their own reach. Exploration's primary driver for outsourcing is to refocus their (residual) IT organisation. Over time and after various debates within the Group, most BP managers appear to have agreed that by following a piecemeal approach to sourcing IT services they gain experience that proves invaluable when they decide to consolidate their separate contracts into one large outsourcing deal. Table 4.1 below summarises the two opposite IT

procurement philosophies of BPO and BPX as described in BP documents and explained by BP managers. These two philosophies seem to mark the two extremes of the range of opinion within the Group.

Table 4.1

Outsourcing criteria, reasons and expectations for BPO and BPX

BP Oil	BP Exploration
• Economies of scale can be realised beyond BP's reach. • Costs can be made variable. This is considered particularly relevant to declining legacy systems. • Mature, low-risk, well understood services are better suited to outsourcing. • After compacting and improving a service internally to the point that no further improvement can be achieved in-house. • Where external skills can be accessed cost-effectively. • Functions that BPO do not want to manage.	• Refocusing the residual IT organisation. • Leveraging relationships to access new technologies and expertise. • Accelerating the pace of technology delivery and installation. • Improved cost focus by avoiding infrastructure investment. • Reduced and variable costs.

5. BPX supplier selection

BP Exploration's approach to outsourcing and the supplier selection process they followed are described in a Harvard Business Review article by John Cross (1995), BPX general manager of IT. The process started in 1991 and was thoroughly prepared. Initially they mailed requests for information to 100 vendors including their incumbent suppliers and three internal management buyout bidders. Of these, 65 responded and an initial screening by a team of 20 BPX managers from various departments reduced the list to 16. At that time (June 1992) the Corporate Centre had also decided to downsize their operations and joined the BPX initiative. Each of the 16 candidates was visited by BP staff with a comprehensive list of more than 12 evaluation criteria. The list of 16 suppliers was reduced to 6. Instead of inviting them to tender, BPX and BPCC decided to host a 'BP IT Partnering Workshop'. BP

presented detailed requirements for each service, excluding performance (service level) targets. They did however state their cost targets. These were ambitious enough so that candidate suppliers were deterred from underbidding opportunistically. The suppliers were then asked to negotiate among themselves and to make coherent proposals as to how they would cooperate in order to provide a seamless service for the whole set of IT services required by BP. Five proposals were submitted. The proposal finally accepted involved two site providers (SEMA and SAIC) and a WAN supplier (Syncordia).

BPX wanted the three vendors to form an alliance and enter in contractual agreement with BP as a single party/entity. This would allow BP first to have a single interface both for negotiations and for service management and second to ask for seamless service provision across services and business units. This scheme would have significant operational advantages for BP as opposed to an arrangement in which they would have to manage three suppliers and the inter-relationships among them. The then European Economic Community (European Union now) intervened and the courts did not allow this arrangement to materialise on the grounds of protecting competition.

SEMA has taken over data centre operations across the UK and moved them to their own sites, as well as providing IT services for the BPCC and BPX's London-based headquarters. SAIC (Science Applications International Corporation) manages IT services at BPX's European headquarters in Aberdeen and all the applications. It is noteworthy that most BPX applications are specialised, high technology systems for the study of geological data, drilling sites and so on. Syncordia provides WAN services covering the UK and most BPX overseas sites.

SAIC and SEMA are responsible for local site services such as desktops, applications and LANs. Syncordia (BT's subsidiary) has undertaken wide area network services including telephones, video-conferencing and data communications. The main sites

served are the London headquarters hosting the Corporate Centre, BPX's headquarters at Stockley Park, Sunbury and Aberdeen.

6. The transition to outsourcing

Despite the stringent provisions of the law, BPX bypassed the transfer of undertaking (protection of employment) (TUPE) regulations of 1981 by firing and paying redundancy fees to staff prior to outsourcing but informally agreed with the suppliers that they would give them the option of applying for a job to them. Several BP employees were subsequently employed by one of the suppliers. Therefore the outsourcing contract did not include any provisions for personnel transfer.

The contract provided for the transfer of some assets, software licences and prior contracts. Generally, however, they tried to avoid asset transfer if it involved many complications such as licences, leases and many third-party contracts. A senior BPX manager felt that any asset purchases by the supplier at the beginning would be charged back to the customer at a later stage indirectly via the pricing structure. Further, the supplier may prefer to purchase existing assets or retain ownership of new ones, in order to improve its bargaining position in future negotiations and at contract renewal. In the case of BPX, desktop hardware (mainly PCs and printers) was sold to the supplier and leased back. At contract termination this equipment reverts to BP after payment of its residual value (which forms part of the stranding costs - see section 7).

Project teams from the suppliers went to BPX sites during call-off contract negotiations to collect detailed information about each service. BPX did not have to collect and present this information. At this stage the negotiations generated a lot of friction and nearly fell through as the suppliers gradually realised the complexity of services.

Air BP and SEMA agreed on a transitional period during which the two parties would collaborate closely in service provision while SEMA acquired and developed the business- and systems-specific knowledge required. For this purpose, key IT staff were transferred to SEMA. BP's intention was to concentrate on business results rather than the process of service delivery as soon as this transitional period ended. This practice corroborates the theoretical discussion on the role of knowledge in economic organisation that has been elaborated in section 6 of chapter 2.

7. The structure of contracts

The framework agreement is entered into by BP and each supplier. It covers the fundamental principles of the exchange, describes the way business is to be done and the way specific goals are to be set. Moreover, it sets policies for intellectual property and data protection. An open book policy is established according to which BP retains the right to access and audit the supplier's cost information. The issues of measurement examined in chapter 2 can inform such observations. The framework agreement also specifies how disputes are to be dealt with and makes provisions for service level reviews (technical aspect) and renegotiation of payments (commercial aspect). Cost targets are reviewed and negotiated annually but service level targets rarely. Two categories of costs are defined. Direct costs relating to resources used directly for service provision such as hardware, software, staff and subcontractors. Indirect costs are management and other overheads. The agreement also allows BP to benchmark the suppliers' costs against prices in the market. Suppliers are required to hire the best-in-class subcontractors whenever subcontracting is required.

The framework agreement also makes general provisions for contract termination. Standards and procedures are specified to ensure uninterrupted service and initial support to the succeeding provider. These are relatively few and are not defined in extensive detail. Performance expectations are detailed in the call-off contracts. A

senior BPX manager feels confident that BP will be able to change supplier promptly if needed. In his opinion, if it is possible to switch supplier it is equally easy to bring a service or a set of services back in house. Exit provisions include the payment of so called 'stranding costs' from BP. Stranding costs include the amortisation of capital investments and the salaries of staff the employment of which the supplier has to maintain. Stranding costs in this case can be seen as a measure of the specificity of the supplier's investment.

The framework agreement also describes an escalation process which is to be followed when problems persist or major disputes arise. The process gradually involves more senior people from the two organisations who attempt to reach an agreement before the problem is referred to the next level of escalation or to arbitration or to even more severe legal measures. The following managers from the part of SEMA are involved at increasing levels of the escalation process: site manager, account manager, area manager, production director and managing director. From the part of BP it is the partner resource manager who makes all these contacts but in severe cases more senior people within the BP organisation may have to exercise their power of authority. The contract stipulates with precision the conditions under which each level is invoked. Overall, it has proved to be a useful and effective measure, particularly where the performance of third-party subcontractors is concerned. When their performance is not up to standard and BP managers cannot bring about any correction, SEMA are notified and their bargaining power over third parties is such that any problems are promptly resolved. When SEMA is about to renew their contract with a third party, BP have the right to intervene and make specific requests, within certain limits, but they seldom make use of this right. These provisions demonstrate the significant relationality of the contract (see section 2.1 in chapter 2).

Call-off contracts are negotiated separately by each business unit at each site with each supplier and may include several services. They are subject to the provisions of the framework agreement but may identify particular amendments that suit

individual circumstances. Call-off contracts specify the following: charging and billing, the service level agreement (SLA), site management organisation, the scope of the contract, plans for maintaining and improving quality, the particular grief points and incentives applied, confidentiality clauses, transfer of hardware, software, licences, documentation and continuing contracts. The case of WAN services is an exception as it was treated as one business/one site and was specified in one call-off contract with Syncordia. The cost is shared by BP businesses/sites.

BPX call-off contracts provide for a regularly renegotiated cost payment, plus a fixed percentage margin. They also include a formula for yearly cost reduction according to changes in market prices, the economies achieved by the supplier and other economic parameters. BPX retains the right to benchmark suppliers' services and costs against competing vendors in the market. In the event of a discrepancy, the supplier has to either match the cost or justify it. Cost-plus contracting of this kind provides no incentives (chapter 2, section 4). The benchmarking rights are an attempt to remedy this problem by restoring some competition and thus acquiring comparative performance information.

Detailed service level requirements are specified for each service. For example, the main services included in BPCC's call-off contract are data centre operations, PC and workstation support, help desk, LAN support and applications support. Some call-off contracts also specify the amount of effort required for each service in terms of person-hours or number of staff. Costs for applications support and maintenance are fixed and expressed in terms of full-time employment equivalents of person-hours.

For some infrastructure support services, the continuous improvement of service quality is not considered important as long as certain levels of availability, reliability and response are met. For more involved services (e.g. business applications) quality improvement becomes a major issue. Quantitative service targets within the contracted SLAs do not always prove satisfactory in practice

because they turn out not to cover all aspects of the service. Key-user satisfaction surveys complement discrete targets and are also part of the contract and of the incentive payments. The contracts specify no more than '*simple questionnaires on service quality [to] be distributed to a small number of key users and if their responses are considered satisfactory a payment of 1% [to] be made to the supplier*'[7]. These user satisfaction surveys are administered quarterly and for the suppliers to get the full 1% they must achieve satisfactory results throughout the year. Air BP managers felt however that user satisfaction surveys were not properly addressed in the contract and the feedback collected in practice was inadequate to assess service quality. Some managers felt that BP was altogether too hands-off but perhaps this might have been a natural reaction to the transition from a very hands-on environment to contract management. These issues point to the discussion of information, measurement problems and incentive structure in chapter 2.

Prior to negotiating the call-off contract, some BP businesses benchmarked their in-house services in order to set service levels and cost targets. The supplier was then given the chance to verify BP's benchmarks. Service levels and costs were thereafter agreed upon the assumption that services would remain the same. This assumption may not hold for certain services over an extended period of time. As for new services or for revised service level agreements, joint negotiations on a zero-basis with the supplier have been carried out.

Work not required by the contract(s) or not achievable within the resource levels specified is classified as 'out of scope'. Examples of out of scope services are user requests or system amendment requests (SARs) for applications. Small projects are carried out and charged without further negotiations but for major work either a separate contract has to be negotiated and entered into or the existing SLAs must be renegotiated and changed. This provision allows the suppliers to look forward to

---

[7] Air BP call-off contract with SEMA.

additional business while creating an additional and potentially uncontrollable source of expenditure for BP. The SEMA account manager mentioned that in one year they had 66 cases of small projects in addition to the contract and they were looking forward to more such work. He estimates that 10% of the work they perform is not stipulated in the SLAs but is initiated by the users in the course of contract execution. Such is the nature of contract incompleteness (chapter 2). Contracts are necessarily incomplete, particularly in IT where technologies and business needs change regularly and where user requirements cannot be easily and accurately elicited in advance.

The approach BP wanted to adopt in specifying service level agreements, was to focus on business requirements for each service rather than the technical characteristics and the way services are provided. In practice however, this was rarely possible. Whenever there is a transfer of responsibility of existing operations or applications, the contract specifies in detail, the hardware and software used and the amount of person hours needed. While the rationale initially was to buy the technical expertise of the suppliers and let them decide the best way to provide a service, inevitably, existing systems have to be described in the contract. Moreover, in order to control costs, in some cases BP determine the human resources required. The managers interviewed recognised that specifying the resources to be used in the contract involves a trade-off between cost control and shifting responsibility for service levels. The trade-offs between contracting for input resources and contracting for output are well documented in the theory (chapter 2).

The process for devising or revising performance metrics and targets involves an initial consultative discussion with the users, the partner resource manager (PRM, see section 7.2) and the supplier. Then technical staff are called to assess technical details. Finally, commercial (incentive and financial) issues are dealt with. It is a formal process because if it involves change in the scope of a service, further negotiations may have to be entered into and formal documents must be produced and signed by the appropriate managers from both parties.

7.1. Pricing and Incentives

A fundamental element of the incentive and monitoring system specified in call-off contracts, is the grief point system which varies between the contracts of different sites or business units. When suppliers do not meet the performance targets set out in service level agreements, BP collect points according to detailed rules taking into account the extent of non-performance and the criticality of the service. These points are accumulated over rolling periods of time and complex monetary reward or penalty rules are subsequently applied. According to BP, the purpose of grief points is to maximise flexibility by emphasising that the goal is to get the job done rather than penalise the supplier for every single instance of bad performance which might well be due to unpredictable factors. Any failure to meet a given target could otherwise signal breach of contract which could in turn lead to costly disputes and perhaps litigation which would not only damage the relationship but finally wreck IT services. This is clearly undesirable and the grief point system is instituted in recognition of the uncertainty involved and is designed to allow uninterrupted service provision and smooth trade relationships as long as performance is within some acceptable tolerance limits. The penalty system is comprehensive and forgives isolated exceptions, penalises persistent problems and allows BP to terminate the contract unilaterally when performance falls very low.

Performance indicators for each service include service level metrics and activity measures. The former are distinguished by the fact that a given performance target is specified for each service level metric. Performance targets can be either absolute quantities (e.g. 'the European network is allowed to suffer no more than 5 outages per month') or percentages (e.g. '50% of help desk calls must be fixed in less than 5 minutes'). Service level metrics are classified as either critical or non-critical and have a different 'grief point value'. Service criticality is judged and negotiated on the basis of the importance of service in terms of business performance. For

example, database availability at the end of each quarter is more critical than database availability at other periods of time.

Activity measures do not convey a performance target but are additionally reported on a regular basis in order to collect richer performance information. An example of activity measure is a monthly report on the total number of videoconferences requested, held or cancelled. This information is used to qualify the two related service level metrics, namely, that 98% of conferences must run without problem or delay (critical) and that no more than 1 VIP conference to run with a problem or delay (non-critical) per month. Demand information provided by the report (activity measure) can help infer whether the service levels actually delivered (as measured by the two service level metrics) were easily achieved or whether any problems were justified. The discussion of information and measurement in section 6 of chapter 2 is pertinent in this respect.

Service level metrics and activity measures are measured and reported monthly, weekly or daily. Some are only reported by exception (i.e. only when deviations or problems occur). Certain metrics are characterised as 'mystery shopper'. These are monitored occasionally. Such metrics are either easy to achieve and hence rigorous monitoring is not necessary (e.g. 'provide telephone directory to users on the same day of the request') or their measurement is involved and complex and hence cannot be frequently carried out (e.g. for help desk services, 'contact a sample of users to confirm problems are handled satisfactorily').

Metrics are recorded on manual logs or by specially written software tools, wherever possible. BP and the three suppliers agreed on a limited 'honeymoon' period in the beginning of contract execution, during which grief points would be calculated but not enforced in order to test all the arrangements and address any teething problems. In the early period of the outsourcing relationship some gaps were found in the service level agreements. The supplier agreed to incorporate additional metrics and by the time the case study was taking place they were

considered complete. Over time, certain metric targets have been renegotiated but not the metrics themselves.

It is the supplier who has responsibility for keeping track of metrics and reporting them as required by the SLA. Moreover, a weekly copy of the help desk log is handed to BP. BP retains the right to audit the validity of the suppliers' reports but they seldom make use of it. At the time of the case study they did not have the resources or the procedures to carry out comprehensive and regular audits. Normally they rely on informal, approximate judgements on the supplier's truthfulness. It was felt that any misrepresentation of performance would be detected or inferred by the existing managerial structures and would subsequently trigger a full audit, if necessary. This example provides an interesting qualification to the theoretical arguments on opportunism and misrepresentation of information. Concern was also raised over making the structure and content of performance reports more efficient and useful for service management.

The grief point system as applied to the service level agreements with SEMA for services delivered to Oil UK and Oil Europe combined, is based on a monthly monitoring and reporting cycle. As far as critical metrics are concerned, if all the targets are achieved, then a point is awarded to SEMA. If SEMA fail to achieve between one and five of the targets, a point is awarded to BP. If SEMA fail to achieve more than five of the targets then two points are awarded to BP. For non-critical metrics, consecutive failure on the same non-critical metric for 3 months will result in a point being awarded to BP and the counter for this metric being reset. If SEMA collect three points, the counters for the critical metrics are reset for both SEMA and BP. If BP collect three points then 10% of the quarterly management fee is paid to BP and the counters are reset for both SEMA and BP. If BP are awarded two payments in a rolling twelve-month period, BP can review the option to terminate the contract on grounds of breach. If BP elect not to terminate the contract then they receive a payment of 10% of the quarterly management fee and BP forego the right to review the option of contract termination until two

further payments are awarded to BP. These rules are depicted in the flow chart of Figure 4.2.

A number of activities have not been assigned a measured target but have been assigned a qualitative expectation. If it becomes apparent that this expectation is not being met, then BP and SEMA will agree a way to include these in the grief point mechanism.

The grief point system contains an ambiguity related to the assignment of points to BP. If BP has collected 2 points and during the following month all critical targets are met then BP's points are reset to zero. However, it is possible that during the same month BP is awarded further points due to the supplier's consistent underperformance in non-critical metrics. In such cases it is not clear what action is taken (financial penalty or reset points). BP managers maintain that since it is their responsibility to administer the grief point system, they make judgements on a case by case basis and try to address any problems in cooperation with the supplier.

The overall effects of the grief point system appear to be ambiguous. Its mere existence makes users more confident that their expectations will be met. A partner resource manager, however, had some reservations regarding its effectiveness. Gaps in the specification of metrics and in the operation of the grief point mechanism are unavoidable. BP managers felt that the supplier has been able to take advantage of such gaps in small and incremental steps, at the expense of the quality and cost of service to BP. At the time of the case study, BP managers were seeking supplementary instruments to guard against such behaviour of the supplier and to provide further incentive for innovation and service improvement. It appears that opportunism is manifest in taking advantage of gaps in the contract (contract incompleteness) rather than in misrepresenting asymmetrically available (albeit measurable) information (see chapter 2).

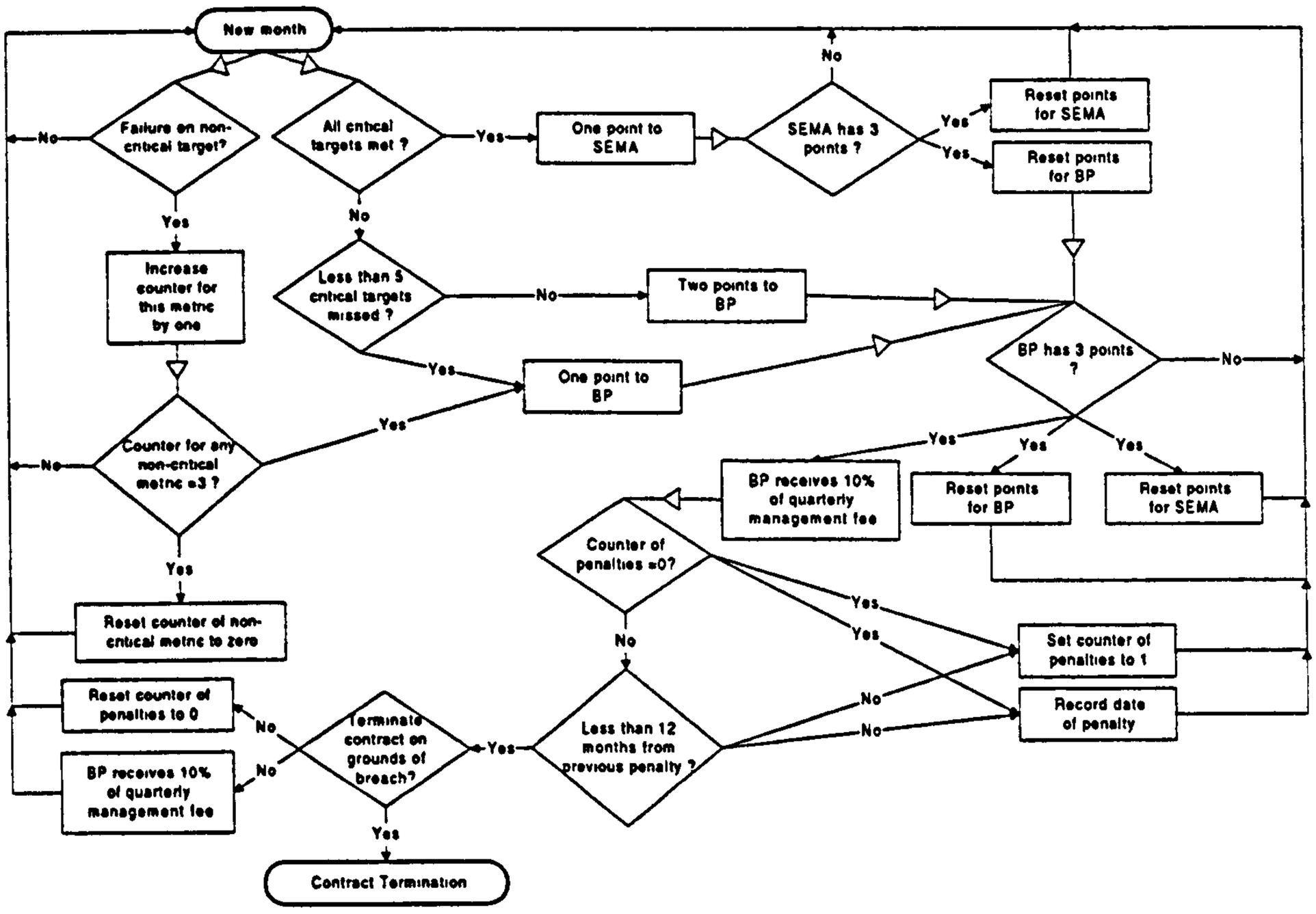


Figure 4.2

The grief point system of the BPO/SEMA call off contract

An incentive system that does not reward good performance but penalises consistently bad performance instead, while protecting against uncontrollable contingencies.

User surveys are a service level measurement instrument which has not been employed in the BPO/SEMA contract but they were planning a way to develop and introduce them formally in the contract.

The incentive system stipulated by the Air BP contract has also been examined. For 1994 the incentive payments (target) were to be calculated on a subset of selected items (including permanent and contract staff, systems maintenance, two applications and a data centre) from the total cost target for the contract. Incentive percentages and the target were to be renegotiated every year. For 1994, rewards were to be up to 2% of the incentive target and penalties up to 1% of the incentive target. On top of that there was to be an additional payment of 1% of the incentive target if the perception of designated key users (derived by means of a survey) is that the service received is acceptable. The supplier's incentive would depend on solving all problems even those not directly caused by them or under their own

responsibility. This is justified "in the interest of making agreement of any incidents easier"[8]. By foreclosing the allocation of responsibility, this contract, clause mitigates potential costs of coordination, haggling and influence activities (see chapter 2, mainly section 4).

The incentive system aggregates the supplier's performance over five major services, each of which specifies targets calculated on a number of detailed requirements. The five services are: database availability, front end processors, telex and fax system, availability of invoices and statements, priority 1 incidents. All these services are measured monthly or averaged over monthly periods. For each service intervals of performance are specified for awarding -1, 0, 1 or 2 percent of the incentive target to the supplier. The percentages are averaged over quarters and then over the year and the incentive payment is settled at the end of the year. As for the measurement of user perceptions the contract specifies 8 key users to which a simple questionnaire is sent every quarter. If the supplier's service is judged satisfactory throughout the year, another 1% of the incentive target is awarded.

The incentive systems of Air and Oil are very different, despite the fact that they are under the same framework agreement with the same supplier. Oil's system penalises the supplier in cases of unacceptable performance and does not provide any positive rewards. Air's contract, in contrast, emphasises monetary rewards for good performance while including a smaller penalty for non-performance. Following the theory of chapter 2, this could be explained by the different types of services included in Oil's and Air's contracts. The services included in the BPO contract are comparatively more standard: invoice management, video-conferencing, telex, fax, voice, network operations, network installations, help desk, office moves, commissioning, installation and acquisitions, PC support, e-mail administration and support. Air's contract is for the operation, maintenance and support of business applications which are by definition specific to the

---

[8] Air BP Limited / SEMA Call-Off Contract, November 1993.

organisation. Applications use a total of approximately 16 databases which are also included in the contract.

Another difference between the incentive systems of Air and Oil is their intensity. Air's contract awards the supplier between 1% and 3% of part of the cost target for the year and/or charges the supplier up to 1% of the same cost target, up to three times a year. Oil's contract may charge the supplier for up to 10% of the quarterly management fee for up to two times in a rolling 12 month period. It is difficult to make an exact comparison of the two systems as they are calculated in different ways. It appears however that the Oil contract is more strict and punishing while Air's is more encouraging. This is perhaps due to the fact that the former contract involves standard commodity services which should create fewer uncertainties to the supplier, as compared to the more business-specific and complex application services in the latter. Service level requirements are more easily specified, administered and relied upon in the case of standard and mature services. Penalising incentives are thus expected to be fairly efficacious. In contrast, business-specific applications are not as well defined and involve significant uncertainty. In this case, penalising incentives would not be sustainable and rewards are more appropriate. See also chapter 2, section 4.1, paragraph 6.

The part time Partner Resource Manager of Air BP mentioned that the user survey specified in the contract has not worked very well. Nevertheless, this is not considered a major problem as long as costs are acceptable and critical services are delivered. It was not clear however how they would react if major problems arose.

An incentive system similar to OEU but even more intense is implemented with BP Finance. In particular, SEMA gets six points for not achieving critical targets and one point for non-critical targets. They have to pay a penalty when they have collected a number of points but they also receive a reward for achieving the targets. This incentive pattern can be explained if the systems and services at BP Finance are relatively standard (so that severe penalties can be acceptable by the

supplier) while highly critical to the performance of the business (so that severe penalties are also justified). To the extent that uncertainty is associated with certain applications, positive rewards for good performance compensate losses due to 'bad luck' in other instances.

In January 1995 BP Oil started to negotiate a new set of quality incentives to be introduced in the revised contract between BP and SEMA for the Corporate Centre in London. According to their draft plans, part of SEMA's profit margin would be contingent upon the fulfilment of five categories of quality criteria. It was proposed that each category would bear equal weight and the rating of SEMA's performance against these criteria would be at BP's descretion. The five categories of quality criteria were first, criteria related to the supplier's responsiveness, i.e. it's ability to suggest innovative ideas for service improvement and cost reduction. Second, criteria pertaining to the collaboration and information exchange between BP and supplier staff in formal and informal structures. Third, criteria related to the usage, management and review of service level metrics and targets. Fourth, criteria encouraging new services and adaptations to existing ones according to business changes. The fifth criterion is user perception of service as derived from 'customer responsiveness surveys'. These criteria are clearly difficult to measure and verify objectively and consistently but BP wanted to introduce them regardless, in response to problems faced during the first couple of years in the initial contract.

For example, innovations in service delivery made by one team or supplier was occasionally withheld from others in other sites since they were competing for a larger share of BP's business. As a result, the proposed arrangements include incentives for increased communication and diffusion of new ideas. In other instances the suppliers were sometimes reluctant to introduce changes in the services and saw such cases as opportunities for additional business at a high premium. The proposed incentives prescribe ways for systematically taking into account requirements for change. The problem is, of course, that the proposed criteria are rather abstract and difficult to measure and verify objectively. Agreeing

on and implementing such provisions could therefore be difficult. This appears to be an attempt to address some of the effects of contractual incompleteness by instituting a process of subjective assessment. Further details on this attempt have not been made available.

BPX contracts determine a profit margin for the supplier comprising a small fixed part (up to 25%) which is a percentage of total costs and a large (75%) variable incentive part, geared to the achievement of cost-reducing targets. A senior BPX manager acknowledges that suppliers have their own incentives to perform well, beyond the formal reward system stipulated in the contract. The reputation and credibility suppliers try to build up in the marketplace constitute significant performance drivers. BP is a big customer and the potential for more business within the group is an incentive to demonstrate good performance (see section 7, chapter 2).

Other BP managers however repeatedly emphasised that while they were generally satisfied with the suppliers, they noticed that whenever possible they (suppliers) try to take advantage of minor gaps or ambiguous parts of the contract. Taken individually, such instances have negligible impact but if compounded they constitute a significant benefit to the supplier at the expense of BP. Such a pattern of opportunism features prominently in the academic literature (chapter 2).

7.2. Site organisation

The organisation of site management is specified in service level agreements. A typical site management organisation consists of a Partner Resource Manager (PRM), a technical and a business consultant from BP. From the supplier's part there is an Applications Support Manager (ASM), an Account Manager (AM), a Site Service Delivery Manager (SSDM) and a number of other full- or part-time staff for roles such as personnel, legal and contract management, production services,

technical services, quality, financial control. The exact organisation on each site varies depending on the types of services delivered and the total size of operations.

The business consultant's role is close to the business and the users. His or her responsibilities include understanding business strategy, aligning systems with business strategy, helping users to develop their information strategy, monitoring the delivery of systems and ensuring the latter add value. The technology consultant is placed closer to the supplier and to the partner resource manager. His or her role is to contribute to the development of the business unit's technology architecture and to the integration of systems and to advise on how best to take advantage of emerging technologies.

The role of the partner resource manager is often filled on a part time basis. This is the case, for example, with Air and Oil/Europe. He/she is responsible for day-to-day service delivery and for handling the escalation process. His/her responsibilities also include verification of supplier's monthly reports and SLA amendments in consultation with the customers.

The performance of the PRM depends on the individual and has repercussions for the ongoing relationship between the parties. One SEMA account manager reported a case where their reputation in BP suffered because they had to have many contact points with the customers while one PRM was not fulfilling his duties. He believes that the role of PRM does not provide adequate incentives to perform well (but he didn't define good performance). He also finds BP's supervision of SEMA's operations too close while he does not consider it necessary.

The Applications Support Manager and the Site Service Delivery Manager are assigned to different sets of services supplied to a particular site, typically applications support and desktop support respectively. Their titles and responsibilities may differ from site to site depending on the services being provided. In general, their role is to manage day-to-day delivery of some given

services. Different services (e.g. applications and mainframe operations) require different management processes and skills and are therefore assigned to different managers. The account and site managers are the main contact points of the supplier organisation for BP. Both interact mainly with the Partner Resource Manager. Their primary responsibility is to manage and coordinate SEMA's contractual obligations and to negotiate contract amendments and scope changes. The account manager is based at SEMA's headquarters while the site manager is more hands-on and located on-site.

The framework agreement specifies the role of the outsourcing review board (ORB). This is a joint committee which meets quarterly to discuss dispute resolution, persistent problems and changes in the scope of the contracts. The ORB is chaired by a senior manager from the BP group and participants include the heads of Information Systems of the BP businesses for which the Board convenes, the respective account manager and a senior manager from the supplier. This is a further element of relationality in BP's contracts (chapter 2, section 2.1). The outsourcing review board was not considered to function effectively and fulfil its intended goals. It is considered to lack focus, because first, its members are frequently replaced and therefore it is difficult to develop and maintain continuity in approaches and accord and second, participation has been irregular and inconsistent.

A monthly service review meeting is held to review service level reports on a service by service basis, to assess the status of grief points and to action decisions. The meeting consists of the site manager, the area manager, the account manager and, less often, the production manager of SEMA and is chaired by the partner resource manager from BP. The same participants also hold occasional meetings to review major or persistent problems and devise measures and policies to avoid similar hazards in the future. According to the SEMA account manager, the service review meetings work well, in contrast to the outsourcing review board.

IT managers at Oil perceive their role as intermediators between suppliers and end users. Service Level Agreements (SLAs) are instrumental in this structure. End users have business problems and requirements and the systems (IT) group is called to deliver business solutions. Their role is to translate these needs into technical requirements and request them through service level agreements. Because SLAs are necessarily restricted, the IT group also has to monitor the performance of suppliers and manage the expectations of the customers. Rationalisation of user requirements (what BP call demand management) is seen by IT managers as key part of their role, in addition to contract management and performance monitoring. The IT group is therefore not a mere 'broker' or 'clearing house' as suggested in the literature (Loh and Venkatraman, 1995) but a proactive entity adding value in the process of service provision.

From the point of view of the customers, services are provided by the IT group who in turn have subcontracted them to the three suppliers. A problem that has risen in the Hemel Hampstead headquarters, is that there are many groups of customers on the same site and therefore different contracts and different service and partner resource managers with whom SEMA has to interact. This created a coordination problem for SEMA and at the time of the case study they were trying to simplify the interface and create a single contact point between the two organisations.

8. Further issues of concern

The outsourcing bandwagon

BP Exploration initially outsourced only infrastructure services and not application development and support. Soon however they started writing call-off contracts for application services with the same suppliers. The pattern of a bandwagon effect appears to emerge. BPX initiated the outsourcing exercise for infrastructure services (mainly wide and local area telecommunications, help desks and desktop support). Other businesses of the BP Group soon joined in the same framework agreement with the same suppliers and negotiated additional call-off contracts and

service level agreements. The outsourcing epidemic spread not only across business units but also across the range of services. Applications (i.e. business-specific IT services) were gradually handed over to the same suppliers. Several plausible explanations for this could be postulated. Either BP businesses were happy with the level of service and cooperation achieved and were confident to broaden their contracts, or the initial leak of expertise started an avalanche in which it became increasingly difficult and costly to maintain any in-house IT teams with outsourcing becoming a self-justifying option, or the outsourcing trend was part of the general restructuring plans orchestrated by the Group's top management as part of 'Project 1990'.

Market competition versus transaction costs

A related issue of concern for BP IS managers was whether it is worth introducing competition in their IT procurement schemes by allowing more than one vendor to supply the same service, perhaps in different sites or business units, thus creating a 'stick-and-carrot' incentive for higher quality and lower costs.

BP recognised that the processes of supplier selection, tenders, negotiation and contracting from scratch are very costly. This is one of the reasons why they have decided to hold on to the set of three suppliers they initially selected, maintain the framework agreement and use call-off contracts whenever they need to purchase a product or service, as long as they are reasonably satisfied with the service received. In this way the process of supplier selection, a large part of contract negotiations (the part corresponding to the framework agreement) and the respective transaction costs are avoided. Both applications (business-specific) and more commodity services are included in the same contracting approach.

The maturity both of the IT industry and of BP with given technologies is considered significant in the decision to outsource or not and how to structure the outsourcing relationship. Moreover, long term contracts are considered necessary because the supplier needs a guarantee that he will be doing a fair amount of

business to justify the necessary resource commitment. This guarantee of long term exchange is also an implicit incentive. But long term contracts raise questions of flexibility from BP's point of view.

While BP consider their relationships with their suppliers to be a partnership, they have insisted in making their contracts as comprehensive as possible. They acknowledge that the notion of partnership, however elusive, is useful given the complexity, uncertainty and the unavoidable incompleteness of contracts involved.

The impact of the decision of the European Court was significant for what BPX was trying to achieve. The initial plan was to enter a single contract with all three suppliers and thus delegate to them any problems of coordination amongst them while ensuring a single interface for each user site. This ideal would implement the BPX vision of focusing on supply management and providing a truly seamless service. Now BP has not avoided the burden of managing three separate contracts, the coordination between them and any conflicts.

More obstacles to the totally seamless, end-to-end service vision arose when they started writing service level agreements. For many services it was impossible to specify only business requirements. Particular systems and technologies also had to be stipulated, thus transferring part of the responsibility for choosing the technologies to BP. In order to control the costs, staff numbers and skills were also specified, thus increasing BP's involvement in the details of service delivery.

Measured and true service quality

A widely held belief among BP managers was that service level agreements and service targets ought to be reviewed every year in order to take into account aspects of services previously neglected and to expand service targets to include more elusive requirements such as user satisfaction, value added and user documentation. The primary aim during the first year of the contracts was to get the services running smoothly. They anticipated however that once this is achieved,

further improvements could be sought. The problem they observed was that the contractors inevitably stick to the letter of the contract and provide only the requirements explicitly specified. However, the quality of complex, business-specific services is multifaceted and difficult to define on paper. Consequently the real satisfaction with service provision might be different to what formal performance metrics show.

The outsourcing of desktop support at the corporate centre seems to have created some discomfort to end users who were not interacting any more with the people they knew but with new personnel who might or might not adapt to the culture of the office and the idiosyncracies of individual users. Users felt that they lost control over such 'soft' issues, although the contract gives the right to BP to remove particular people from the site when serious problems arise.

BP managers were also concerned with the fact that user satisfaction and other service level measures for a given service may depend on other services. For example, user satisfaction with e-mail depends not only on e-mail support but also on infrastructure and telecommunications operations. This has implications for the allocation of responsibility between contractors. The supplier providing e-mail support should not normally be charged any grief points if e-mail fails because of a telecommunications problem which is the responsibility of another contractor. Identifying and verifying the source of bad performance in a web of interdependent services may be troublesome. However, BP's contracts stipulate that the single supplier coordinating service provision at any site be held responsible for work subcontracted to third parties. For example, because SEMA is responsible for the Stockley Park headquarters of BPX they are also responsible for any problems occurring at this site due to work performed by SAIC, Syncordia or any other third party. The theories of agency reviewed in chapter 2 can recast these concerns in a more informed way. In particular, the problem described here is a typical case of the problem of team organisation, where performance information does not discriminate between team members and the effort each has contributed, and

where overall performance is more than the sum of each contributor's effort (see chapter 2, section 4.1, paragraph 5 and section 6, paragraph 2). Team organisation may lead to phenomena of shirking and free-riding. These theories examine various contractual and organisational alternatives to mitigate such problems. These alternatives include, among others, internal organisation with direct supervision and making team members reciprocal monitors who are motivated with rewards for squealing on each other. In the presence of imperfect information, imperfect verification and divergent interests, all these mechanisms are imperfect and costly haggling (with implications for the cost and quality of service) is expected to result. By making one supplier responsible for all work on a given site, BP passes all these transaction costs to the supplier concerned. Moreover, BP achieves some alignment of that supplier's interests with its own, thus reducing the complexity and severity of the problem.

In terms of user satisfaction the opinion of the SEMA account manager was that in the beginning BP Exploration expectations from SEMA were very high (9/10) and the perception of service delivered lower (5/10). About a year and a half into the contract, BP expectations from SEMA were lower than in the beginning (7/10) and the perception of the service delivered slightly higher than expected (8/10). This pattern reflects the increased hype with which outsourcing was introduced in BPX. Over time, the trading relationship between the two parties adapted to an equilibrium in which users regard SEMA less highly and SEMA in turn manages to satisfy users by slightly exceeding their expectations. As long as this equilibrium is at an acceptable level, serious conflict should not occur. Of course this is the point of view of the account manager of SEMA and BP customers may have other views.

As far as performance metrics in service level agreements are concerned, the concern raised by senior management was that in some instances too many metrics dissipate the focus of overall management and coordination. While detailed and comprehensive metrics may be required, the capacity of management attention and information processing is limited. Therefore some kind of aggregation of metrics is

considered desirable but none has been devised. Nonetheless, such an aggregation is partially fulfilled by the grief point mechanism. Grief points aggregate performance successes and shortfalls in order to calculate incentive rewards and penalties. An overall picture of service levels is thus generated but conveys no background information pertaining to particular services or persistent problems. It is believed, however, that the experience and maturity of BP and the suppliers with a given service as well as the development of a sustainable partnership, allows the introduction of fewer and better targeted metrics, in contrast to services for which little is known and therefore more extensive and detailed information needs to be collected in order to counter the uncertainty involved.

In terms of data security and confidentiality, three types of information are identified. First, information which can be made freely accessible by the suppliers' personnel. Second, confidential information for which every member of staff of the supplier signs a confidentiality agreement. Third, business-critical information which is secret and is maintained in-house and/or is encrypted. Overall, they do not consider data protection to be a problem.

Scope of the contract

An account consultant of SEMA raised the issue of work outside the scope of the contract. The distinction between work included in the scope of the contract and work out of scope is ambiguous and a potential source of friction with BP. When a contingent problem or requirement arises, it is not always clear if it should be part of the contract and therefore be provided by existing resources within the given cost limits or whether it should be considered additional business and paid for separately. Services resulting from business changes are, according to the contract, separate projects. However, it is not clear what constitutes business change. A rough proposal was to incorporate in the contract provisions for additional capacity (staff) to cater for small projects outside the strict letter of the contract.

Related to this is BP's desire that SEMA adds value to the business through IT. For SEMA, added value as well as technological changes would be work additional to the existing contract. SEMA would like to help BP standardise their systems across business units and across sites. This could lead to significant cost reduction. However, customers resist standardisation claiming that their own system has specific features that create business value which is otherwise unattainable.

9. Conclusions

From the various issues pertaining to IT outsourcing that are discussed in chapter 1, the case study emphasises those aspects that are relevant in the context of the theories of chapter 2. Although the organisational background and the reasons for outsourcing are briefly reviewed, they are not the main focus of the case study.

Following transaction cost, agency and property rights theories, the case study emphasises the elements of governance in BP's contracts. Such elements of governance include contractual arrangements that generate incentives, share risks, provide adaptations, distribute information and allocate authority. Broadly, these theories associate different governance arrangements with the characteristics of the service or product being exchanged.

The extent of specificity, uncertainty and information availability with respect to different services were ascertained by discussing and understanding the context of service provision for each service or group of services[9]. The people who can assess an elusive concept such as specificity are the people who know the services best and not the researcher. During the course of the case study, BP managers were

---

[9] Although various measurement instruments have been proposed in the literature (see Rindfleisch and Heide 1997), they tend to be specific to an industry or product, other than the IT services considered here. The development of such instruments would involve a completely different research design, which would not be suitable for capturing the breadth of scope and depth of detail aimed at in this case study.

taught the meaning of concepts such as specificity using examples and metaphors. They were then asked to assess (or debate if there were more than one managers present) the extent of specificity, uncertainty and information availability for different services, in comparative terms. The researcher facilitated the process of assessment or debate. Therefore, there was less need to seek detailed information for all services included in the contracts in order to conduct an 'objective' measurement. The evaluation of service characteristics was done subjectively by the managers who were responsible for setting up and managing the outsourcing relationships.

Since 1991, the BP Group in the UK has outsourced most of its Information Technology services. It is a rich showcase of diverse practices in contracting for IT services. Of course, it does not show all possible approaches. It is also not considered to be a representation of what companies generally do in practice. Finally, it is not an archetype of what should be done in practice. This means that the outcome of the case study (chapter 5) cannot be assumed to generalise to all kinds of IT services, to all possible governance alternatives or to all organisations. However, because of the diversity of BP practices and by virtue of the generality of the theories, the framework can be expected to be more generally applicable[10]. Chapter 6 demonstrates some interesting findings in this respect but, clearly, more research is needed.

The case study examined some perceptions of positive achievements and negative outcomes but it was too early to make firm judgements regarding the success or otherwise of any of the contracts. Any discussion of success or failure in the previous sections, is made with reference to particular governance mechanisms, such as the grief point system or end-user surveys. Besides, the purpose of the case was not to evaluate outsourcing decisions.

---

[10] BP managers expected the S-CAGE framework to be applicable to non-IT procurement as well and were interested in exploring this possibility.

The BP case was used as an instrument to bring economic theory and management practices in IT outsourcing closer together. This was a bi-directional process. First, the case material highlighted a number of important practical issues and generated a further understanding of them, beyond the second-hand knowledge received from the literature (chapter 1). In this way it was possible to maintain a wide reach on the economics literature (i.e. avoid focusing on one or two narrow streams of theory) and to use it effectively by emphasising those parts which were most pertinent to the practical issues raised. Second, the relevant theory was used to analyse the issues identified in practice and to generate structured recommendations. The S-CAGE framework presented in chapter 5 is an organising framework hosting this bi-directional process. It was born through this interactive case study to provide a common language for addressing a common problem from two different perspectives, the practitioner and the theoretical.

Chapter Five

Service Clustering And Governance Establishment (S-CAGE): Bridging Theory and Practice

1. Introduction

The 'Service Clustering and Governance Establishment' (S-CAGE) framework presented in this chapter is the principal contribution of the thesis. It emerges from the managerial insights of chapter 4 acquired through the theoretical lens of chapter 2. All the examples quoted in the rest of the chapter are drawn from the experience at BP. S-CAGE is oriented to both theory and practice. In particular, it is aimed at pulling together a broad range of theoretical contributions, at providing a basis for the study of Information Technology governance and at devising practical guidelines.

The proposed S-CAGE framework is a re-synthesis of the theories examined in chapter 2 and of the empirical evidence from BP analysed in chapter 4. The empirical work at BP was used to bring forth and elaborate the relevance of the economic theories of chapter 2. The purpose of this theoretical background is threefold. First, to assess the empirical findings at BP using the theory as an external criterion and *vice versa* to assess the relevance of the theory in practice. Second, to establish the coherence of the framework on the basis of a single underlying thread. Third, to round out the arguments made by highlighting alternatives not directly observed at BP. Most of the arguments and claims in subsequent sections are illustrated with examples from the BP case.

The theoretical basis of the S-CAGE framework is the microanalytic approach developed in chapter 2, section 8.1. Specifically, it is founded on transaction cost economic arguments supplemented with insights from the theories of (i) agency, (ii) property rights, (iii) measurement, information and knowledge and (iv) culture, trust and reputation effects. Section 2 below builds up the theory behind the basic formulation of the framework. The main lessons learned about the governance of IT from a transaction cost perspective are outlined below.

If service specificity[1] is higher than some threshold level (which may differ from one IT service to another), then external suppliers do not enjoy any advantages in terms of production costs over internal organisation. If service specificity is low, if the company can attain satisfactory resource utilisation and if the minimum efficient size of operations is achievable within the company, then it does not make a difference whether IT service(s) are provided internally or externally. The choice should depend on other considerations.

If complete resource utilisation or the minimum efficient size of operations is not achievable within the company, then external procurement becomes preferable on production cost grounds. If, in contrast, high service specificity is combined with extensive uncertainty and if formal measurement and incentives are not efficacious, internal provision is advantageous on transaction and agency cost grounds. This trade-off requires a comparative assessment of the production and transaction costs involved. Governance structures in the middle range between pure markets and hierarchies provide various contractual and organisational instruments to cope with such conflicting requirements.

There is one crucial decision parameter that is not dealt with exhaustively in this thesis. This is the bureaucratic inefficiency associated with internal service

---

[1] Service specificity is defined here as the extent to which a given IT service involves transaction-specific assets (Williamson 1985).

provision. When such costs are significant, internal organisation may not be superior in some of the cases described above.

As far as scope economies are concerned, the mere existence of common resources is neither a necessary nor a sufficient condition for organising the provision of IT services within the firm (Teece 1980), since the same productive economies can be replicated by contractual agreement between independent companies. Joint production by the same company is preferred from an efficiency perspective when the shared resource is specialised. This is a transaction cost argument and will be further explored later on. Such conditions do appear in the case of IT services, as specialised skills on customer-specific systems and tacit organisational knowledge are always in demand. A counterexample is the computer, which is an indivisible asset that can run a variety of software simultaneously (therefore it can be shared) but in most instances is not specific to any customer. Companies can therefore write appropriate contracts to share one another's computer capacity (e.g. timesharing or reciprocal backup agreements for disaster recovery). In short, those shared resources which are specific are better organised internally (or, at least, under the same governance).

The generally inconclusive considerations about the effect of scale, scope and transaction economies on the efficient organisation of IT, in conjunction with the diversity of actual contracts, point to the need for a closer and more detailed look at the relevant mechanisms. This chapter addresses a wide variety of such mechanisms directly, in an organised manner. The rest of the chapter starts with a brief overview of the proposed S-CAGE framework (section 2). The objective is to provide a concise picture of a number of interrelated issues, which are dealt with in more detail in subsequent sections. One section is then devoted to each of the two major components of the framework: the clustering of IT services into homogenous groups (section 3) and the design of each cluster's efficient governance (section 4). The chapter concludes (section 5) with a recapitulation of the rationale, the scope and the main elements of S-CAGE.

2. Overview of S-CAGE

The provision of IT services is a compound and complex transaction which comprises a multiplicity of different services. For example, wide area network (WAN) provision is a commodity service with significant potential economies of scale as opposed to desktop support, which requires close contact with end users, and knowledge of the specific work environment. Moreover, individual IT services also consist of various tasks which need not be homogenous in terms of their economic properties and the respective implications for efficient organisation. Software development for example requires project management, analysis, requirements specification, design, coding, testing, documentation, installation and training. These tasks are not necessarily carried out by the same provider, whether it is an internal IT department or an external supplier.

A central argument built into the S-CAGE framework is that IT service provision should not be viewed as a unitary exchange. Following the microanalytic research strategy defined in chapter 2 (section 8.1) it is necessary to unpack the bundle of goods and services exchanged within a single transaction (or contract) in order to assess their attributes in detail. For the purposes of this study, the unit of analysis is identified with the single IT service. This is the least divisible unit of IT activity. The next problem, of course, is at what level of refinement to define IT services. As discussed in section 4.1.1, IT services are defined endogenously in the process of devising the governance of IT. The basic argument of transaction cost economics, reinterpreted in this context, dictates that governance structures be aligned with the attributes of each service, assuming that IT services differ in their transaction cost characteristics.

The general question asked is what institutional arrangements maximise the efficiency of IT governance, given the characteristics of the technology, the systems, the services and the markets. Again, the argument made in chapter 2 (section 8.1)

is that the number of actual governance structures is theoretically infinite, ranging on a continuum from pure spot markets to pure hierarchies (Figure 2.2). Figure 5.1 replicates Figure 2.3 and depicts examples of governance structures as commonly known in IT, in the place of the generic governance types of Figure 2.2. Again, the labels commonly used, such as contracting out or outsourcing, specify ranges of possible structures on this continuum rather than unitary instances.

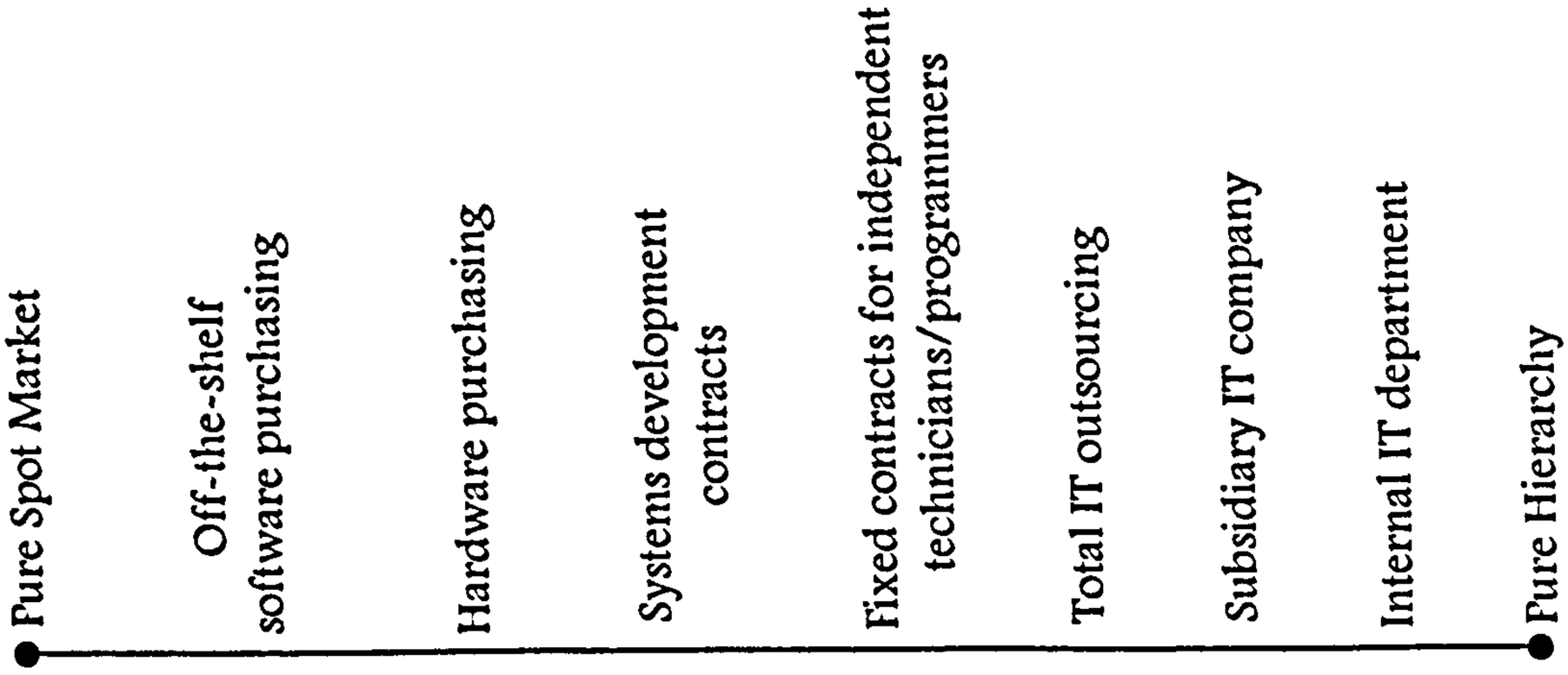


Figure 5.1

The Governance Continuum with indicative IT contracts

Internal IT departments may organise most IT provision in-house but have always had several procurement contracts for software, hardware, maintenance and training, to name but a few. Contrary to some widely held beliefs, total outsourcing contracts are positioned very close to internal organisation. Such transactions, whilst carried out in the market, involve sophisticated organisational arrangements, similar to those found in hierarchical firms. Off-the-shelf purchases of software and hardware are not pure spot market transactions since they usually involve guarantees, training and maintenance contracts. Nevertheless they are comparatively simple exchanges. Systems development contracts are usually positioned around the middle of the governance continuum. They fall into the general category of requirements contracting and involve significant monitoring and supervision by the customer. The hire of independent technicians/programmers or flexible fixed-term contracts is a case of inside

contracting: having someone work on a temporary basis under the internal control of the organisation but with a subcontracting rather than an employment contract.

The underlying argument is for a bottom-up approach to the formation of IT governance, starting with the characteristics of IT services and the components of governance rather than a top-down choice of governance structure.

The argument so far places emphasis on the match between the governance structure and each individual service. This argument alone sheds light on only one part of the problem. Taken at the extreme, each service should be organised separately. However, this is not always desirable or feasible because of technical interdependence, coordination costs, incentive inter-relationships and other inefficiencies. For example Help Desks are services related to many other IT assets and services. The same is also true for desktop support and LAN operations. The quality of many IT services depends on the quality of others. For example, electronic mail, groupware and EDI depend on network operations and support. Similarly, systems implementation depends on analysis and design.

The above examples indicate that the optimisation of governance for each individual service is not adequate for the successful and effective provision of all IT services. Specific governance provisions are needed for the co-ordination of service interdependencies. Service interdependence can be defined in terms of (i) technological indivisibilities and constraints, (ii) the existence of scope economies among services, (iii) the costs of managing many different governance structures (and possibly different suppliers), (iv) the costs related to the monitoring and incentive structures for associated tasks and (v) issues of skills and technical expertise. This argument emphasises synergies and coordination and taken to the extreme dictates the inclusion of all IT services under a single governance.

The efficient governance of IT should match the technological, transactional and agency characteristics both of individual services and of their interrelationships.

This discussion leads to a trade-off between local and global optimisation. Local optimisation addresses the specialisation of governance structures to individual services by favouring separate governance structures for each service. Global optimisation swings towards unifying structures encompassing all services, in order to take advantage of synergies and to minimise the cost of coordinating them.

A simultaneous consideration of local and global optimisation leads to the notion of clustering. IT services are grouped into clusters and each cluster of services is assigned a different governance structure. The governance structure includes the contract plus all those organisational structures and management processes needed to execute it effectively and efficiently. Therefore, there is a one-to-one correspondence between clusters of services and contracts.

In order to strike a balance between local and global optimisation, IT services should be allocated to clusters such that each cluster maximises both the fit between services and specialised governance mechanisms as well as the synergy between services. As shown in Figure 5.2 below, any number of clusters can be adopted, ranging from one cluster for each individual service (the theoretical extreme of local optimisation) to one cluster for all services (the theoretical extreme of global optimisation). Figure 5.2 also depicts the main benefits of one or the other extreme, namely specialisation and synergy respectively. The benefits of specialisation of governance mechanisms to the characteristics of services provide the rationale for a greater number of smaller clusters of services. Taking advantage of technical and organisational synergy requires fewer and larger clusters. Mixed strategies of relatively smaller and larger clusters populate the middle range.

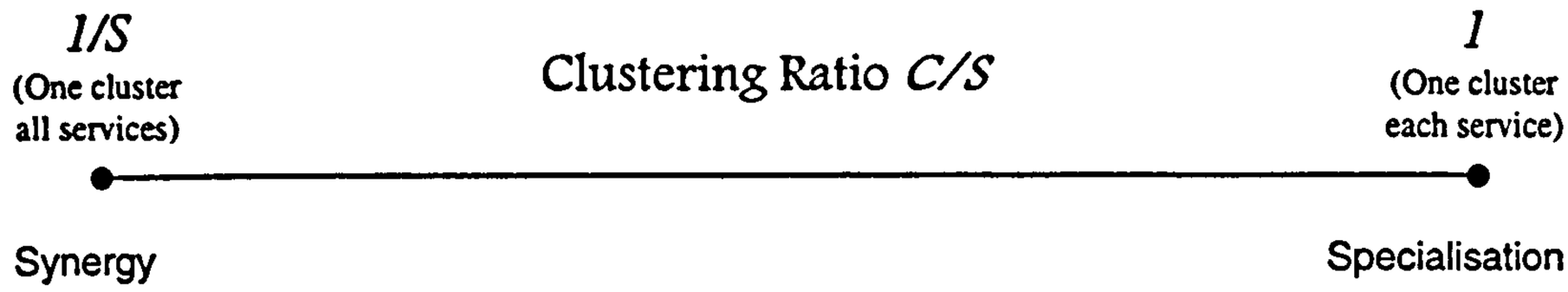


Figure 5.2

The Clustering Continuum and the main benefits of each extreme

The trade-off between taking advantage of synergy among services and capturing the benefits of governance structures specialised to individual services leads to service clustering. The clustering ratio of the number of clusters C over the number of services S ranges between the two theoretical extremes of one cluster including all services ($1/S$) and each service forming a separate cluster (S/S). Synergy and specialisation are the respective benefits.

The ratio of the number of clusters C to the number of services S is a measure of the clustering of IT provision and takes a value in the interval $(\frac{1}{S},1)$. The smaller the clustering ratio the greater the concentration of services to few clusters and the greater the clustering ratio the more fragmented is IT service provision into many service clusters. The two extreme points are theoretical. $C>1$ since, for example, companies never make their own hardware. Total fragmentation is clearly inefficient besides the fact that services are anyway compound and not unitary. Hence $C<S$ and $\frac{C}{S}<1$.

This fundamental trade-off between local and global optimisation in IT governance constitutes the rationale behind the proposed 'Service Clustering and Governance Establishment' (S-CAGE) framework (Figure 5.3) which encompasses both notions of the governance continuum and of the clustering continuum.

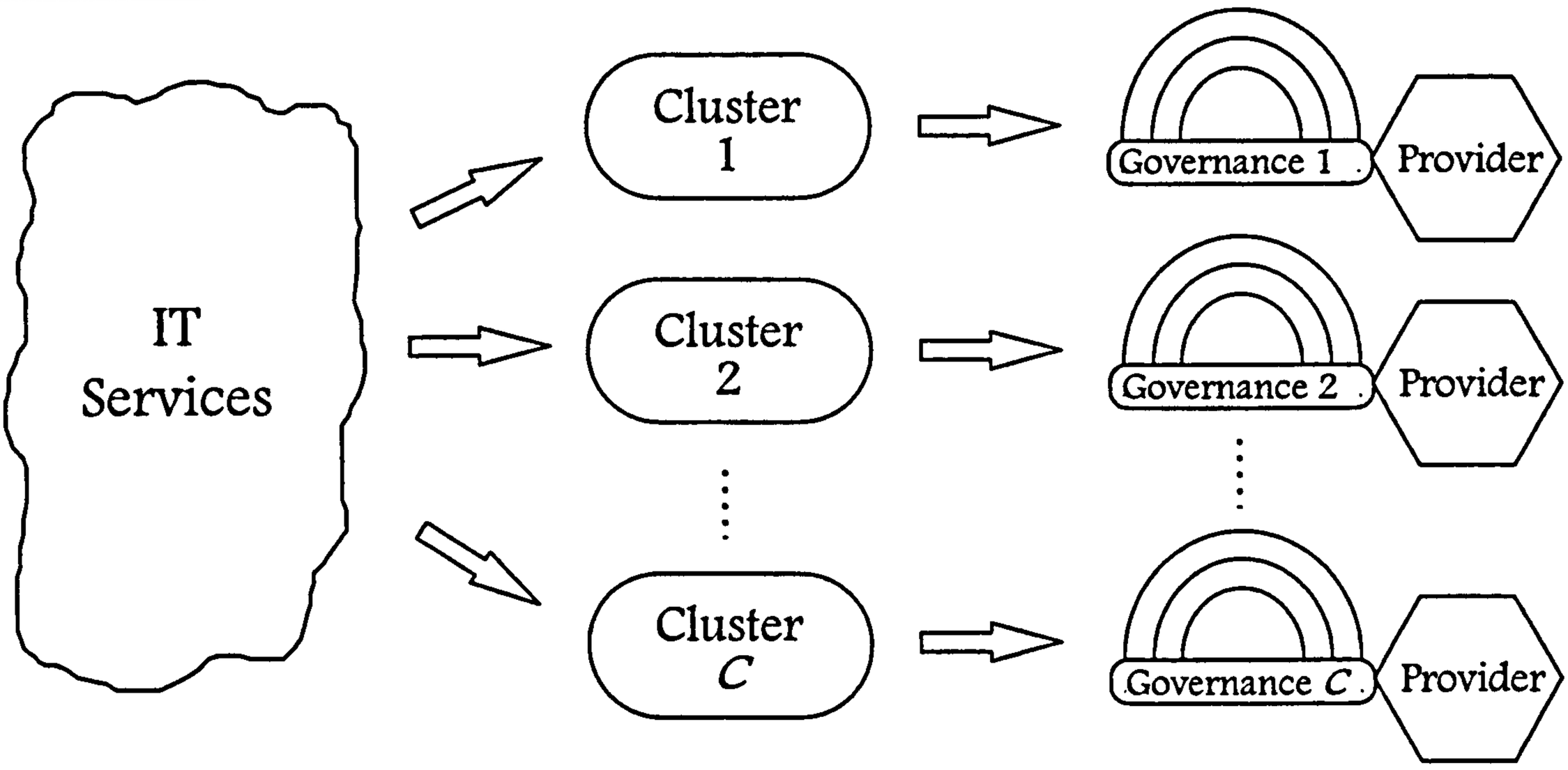


Figure 5.3

IT Service Clustering and Governance Establishment (S-CAGE)

IT services are allocated into clusters. Each cluster is assigned one governance structure (contract). Governance mechanisms are specialised to the services of each cluster. Each contract is entered into with a (possibly different) supplier.

The set of Information Technology services is split into a number of clusters, each of which is assigned a different governance structure (contract) and possibly a different supplier. The greater the number of clusters, the closer the match between the characteristics of individual services and the associated governance (local optimisation) and greater the cost of management and coordination between service clusters. The fewer the clusters, the easier management and coordination between clusters (global optimisation) but certain individual services may be inefficiently organised.

Clustering is not a new idea. Companies have always been bundling IT services and projects in different ways and assigning them to different internal or external providers (see chapter 1, section 6). Moreover, the idea of service clustering is consistent with early arguments that 'you cannot outsource everything' (Huber 1993) and with the notion of 'selective outsourcing' introduced by Lacity and Hirschheim (1995) and Lacity, Willcocks and Feeny (1995). The proposed clustering and governance establishment scheme, however, is more structured and

general, encompassing the whole range of possibilities. It is also substantiated using established theoretical arguments from a range of economic theories.

The S-CAGE framework of Figure 5.3 is a conceptual and not a process model. The arrows do not imply a linear time-line along which services are first clustered and the governance is then set up. As discussed in the rest of this chapter, the clustering strategy depends almost entirely on governance considerations and *vice versa.* Implementing S-CAGE in practice is a different matter altogether. The purpose at this stage is to provide an organised and clear set of conceptual distinctions encapsulating both theoretical insights and empirical considerations.

Plainly, the elements of governance for each cluster will also depend on the choice of service provider (whether it is an internal department or an external supplier), its productive capacities, costs, market reputation and so on, as depicted in Figure 5.3. However, problems pertaining to bidding processes and provider selection are out of the scope of this thesis (see for example Rosenthal et al 1995, Nam et al 1995).

3. Service Clustering

As mentioned earlier, service clustering may result in any pattern on the theoretical continuum between a single cluster encapsulating all services and as many clusters as IT services (each service being organised independently). To decide on the number of clusters and to allocate services into them, a number of criteria need to be taken into account. These are drawn from the elements of governance set out in section 4. One way of deciding on service clustering is to start by considering all services as one cluster and gradually examine reasons for separating different services. This approach emphasises the need to economise on the costs of coordinating different service contracts and to take into consideration complementarities and inter-dependencies among services. The main reasons for separating services are listed below.

1. Other things being equal, services for which specialist suppliers achieve significant economies of scale should be separated and allocated to the most efficient providers. This is the traditional microeconomic argument. Telecommunications and wide-area networks are typical examples. The selection of the three suppliers by BP Exploration reflects such an approach that emphasises specialisation. Telecommunications were assigned to Syncordia (a telecom operator), specialist applications were assigned to SAIC (which has a record of experience in such applications) and commercial applications and services (e.g. data centres and help desks) were assigned to SEMA.

2. Related to the above, if a given service cluster is too diverse for the skills and capabilities of any existing providers, it may have to be split to match what is available in the market. The same applies even when the main candidate is the internal IT department. A company may decide that they cannot sustain skills in certain areas. The respective services may have to be separated. Alternatively, main contractors are often asked to employ third party subcontractors for equipment and services that are not offered by them. For example, BP expects SEMA and SAIC to hire 'best in class' subcontractors and retains the right to intervene in the selection of those third parties. One potential problem with this approach is that the client organisation loses control over the costs and performance of the third party, since it does not control the respective contract.

3. If one service is distinguished in terms of the customer-specific investments it requires, it may be worth separating it in order to reduce dependence on a single provider (in other words, to reduce the amount of appropriable quasi-rents available to any single supplier). If more than one services share the same kind of asset specificity (e.g. skills), they should be clustered together. This derives from the main results of the theory of property rights (see chapter 2, section 5, paragraph 4). BP Oil insisted on keeping the European Systems Project under internal control during development in order to control and protect the business specific knowledge it required. Once it was complete, they outsourced its support

by commercialising it in order to achieve cost savings (see chapter 4, section 3, paragraph 4).

4. Services provided continuously or frequently (e.g. network operations and help desks) may be separated from those provided only once or rarely (such as new systems development) for the same reasons as above. One obvious advantage of separating continuous from one-off services is that it induces competition via the bidding process. The downside is that the process of bidding is costly for all parties and incumbent suppliers usually have an advantage. Otherwise, long term and project-based contracts are markedly different and there are transaction cost efficiencies to be gained from writing separate contracts. See the discussion on frequency as a dimension of transactions in chapter 2, section 3.2, paragraph 2. BP did not make such separation on a large scale. Once the call-off contracts were in place, the incumbent supplier was preferred for any additional or one-off work (see discussions on work out of the scope of the contract in chapter 4). In some respects this is the objective of some outsourcing contracts, to economise on the transaction costs of selecting suppliers and contracting for every task.

5. If greater effort in one service improves the performance of another, then they should be clustered together as the contractor is better motivated by undertaking both. In reverse, if the performance of two services is negatively correlated, they should be assigned to different providers. An example of the first case is software design and coding; design is difficult and time consuming but good design can substantially improve the quality and efficiency of coding. An example of the second case is requirements specification and development: a provider expecting to undertake subsequent development is likely to manipulate requirements so as to maximise his future gains from development (e.g. promote his own capabilities or products while not necessarily serving the customer)[2]. Some BP managers raised the concern that SEMA wanted to 'do things their own way',

---

[2] See mathematical principal-agent theory, chapter 2, section 4.1, paragraph 9.

without necessarily taking into account all of BP's interests (see, for example, chapter 4, section 8, last paragraph).

6. If the performance of each of two services cannot be monitored equally well then under certain conditions they should be kept in separate clusters, because the contractor will place significantly less effort on the service that cannot be as well monitored. This is a result that derives from theory and was also observed at BP. Partner resource managers suggested that one of the drawbacks of the grief point system was that SEMA emphasised the measured targets and ignored others.

Clearly, the above criteria and the benefits of keeping services together may pose conflicting requirements for clustering so that a second-best solution is unavoidable. The resolution of these trade-offs is case-specific. For example, BP Exploration set an *a priori* aim to cluster all services together in order to achieve seamless service provision. On the other hand, BP Oil was more sceptical of this approach and proceeded in a more incremental fashion. In effect, most service clusters (call-off contracts) were determined on the basis of a common site rather than the intrinsic characteristics of services. Chapter 6 discusses some evidence of separating services on the basis of their characteristics.

4. The Governance of Service Clusters

The governance structure of IT service provision can be conceptualised as comprising three interlinked layers, as in Figure 5.4, which is essentially zooming into part of Figure 5.3. The three layers have been derived from the case study at BP as a conceptual classification of the constituent parts of governance. Specifically, contracts and other documents were examined in detail, while various management practices pertaining to the structuring and management of the contracts were repeatedly discussed. A variety of sections and clauses were observed in contracts and a diversity of ways of structuring contractual provisions were encountered. In addition to documented material, the rationale behind

contractual choices and the management processes through which contracts were set up and subsequently managed, were debated with management. The three layers shown in Figure 5.4 are a classification scheme of the elements of governance (which includes contractual provisions and associated management structures and processes) that were observed at BP (see chapter 4, section 2).

As also mentioned in chapters 3, 4 and the beginning of the present chapter, the whole process was constantly informed by theory. The theories reviewed in chapter 2 moderated the process of understanding and classifying empirical evidence, hence they influenced the choice and formulation of the three layers. Incidentally, as it turned out after the S-CAGE framework was formed, in his seminal article Jensen (1983) proposes a three-part taxonomy to characterise organisations, nearly identical to the three layers of Figure 5.4. In that article, Jensen reviews agency and related theories in an attempt to appraise contemporary theoretical developments and suggest further directions in theory development and methodology. Jensen sees organisations as complex systems comprising three subsystems: the performance measurement and evaluation subsystem, the reward and punishment subsystem and the subsystem for partitioning and assigning decision rights among participants in the organisation (p. 325). It becomes obvious below that these subsystems are nearly identical to the three layers of governance proposed here, namely assurance of service delivery, incentives for service level attainment and organisation and management of service provision respectively.

Jensen's scope of enquiry encompasses whole organisations whereas the present analysis is confined to the provision of IT services. However, the provision of IT services requires a set of organisational structures, management processes and contractual provisions (governance structures) which are part of the broader organisation and which are amenable to the same theoretical investigation as whole organisations. Although the three layers of governance proposed in this thesis were derived on the basis of empirical findings at BP, the similarity with

Jensen's classification is not surprising since the same theoretical background informed both schemas.

Each layer of governance is subsequently broken down into three parts that form the headings of the sections further in this chapter. These subsections were devised in the same way as the three main layers (from the contract evidence at BP) and serve to classify the elements comprising each layer as well as to clarify their exposition. The Assurance of Service Delivery layer entails (i) service specification, (ii) service level requirements specification and (iii) service level measurement and verification. The Incentives for Service Level Attainment layer involves implementing and enforcing service level requirements (i.e. the bottom level) with the use of (i) implicit incentive patterns, (ii) contractual reward/penalty devices, and (iii) other organisational and contractual safeguards. The Organisation and Management of Service Provision layer refers to the overall organisational context in which IT service provision is carried out and includes (i) roles and responsibilities, (ii) coordination and communication mechanisms and the (iii) locus and scope of decision making.

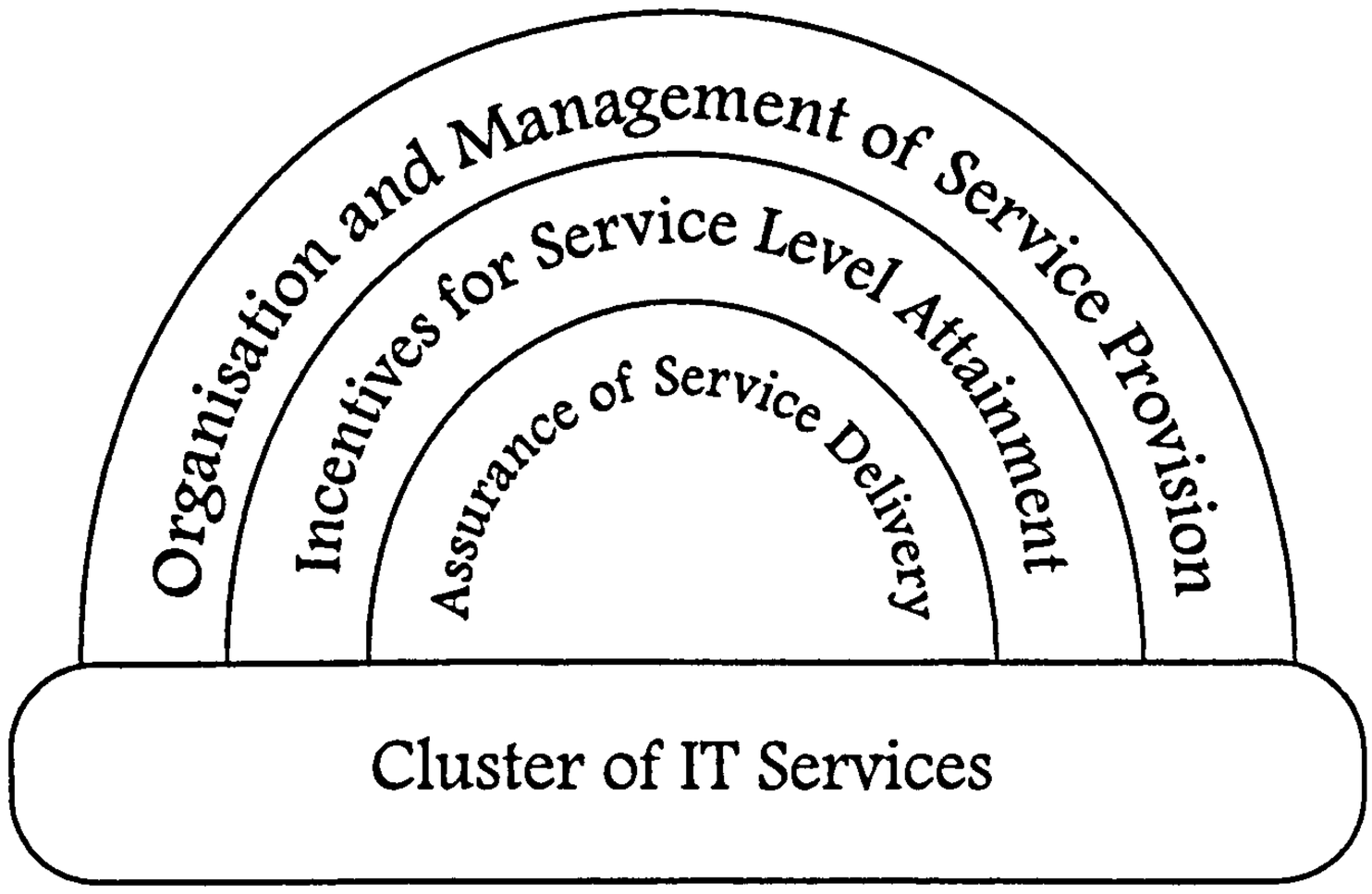


Figure 5.4

The governance of IT service clusters as a layered system.

This layered concept is a framework for organising and categorising the various elements of IT governance. It emphasises both the hierarchical and the enveloping relation between the three layers as well as their direct correspondence with the attributes of the services they attend to[3].

The governance design and decision problems at each layer depend to a greater or lesser extent on elements of other layers. For example, organisational structures (top layer) operate in conjunction with incentive systems (middle layer) which are tightly coupled with performance measurement and verification procedures (bottom layer). The latter are in turn contingent upon the nature of each system or service within the cluster (service specification). The framework of Figure 5.4 is conceptual and the positioning of the three layers represents their hierarchical and enveloping relation to each other and not a process for devising governance structures. Imposing a linear sequence on the various decision problems of each layer as described below is not straightforward and would inhibit the concise and

[3] The resemblance of this diagram to a birdcage, or, in this case, a service cage has inspired the acronym S-CAGE. To some managers at BP it reminded of a sunrise, arguably a more optimistic viewpoint!

integrative exposition aimed at in this chapter. This layered exposition serves as an organising framework of the relevant parameters that need to be taken into account.

The elements of each layer are derived from the contracts studied at BP (chapter 4). The process of derivation was informed by the theories reviewed in chapter 2. Table 5.1 below summarises which parts of these theories are drawn into each layer. The rows of the table correspond to the three layers of governance (Figure 5.4). Each governance layer is broken down into three components, further refining what each layer entails, as mentioned above. Each component corresponds to a subsection in the rest of the chapter. The columns of the table correspond to the headings of sections 3 to 7 in chapter 2, i.e. the theories forming the background of the S-CAGE framework. The cells in the middle of the table indicate what part of each theory informs each (sub-) layer of governance. Grey cells indicate no or negligible relation. More refined references to the relevant theories are made as each layer and sub-layer is explained in the sections that follow.

Layers of governance	Transaction cost economics (ch. 2, section 3)	Agency theory (ch. 2, section 4)	Property rights (ch. 2, section 5)	Information and knowledge (ch. 2, section 6)	Culture and reputation (ch. 2, section 7)
Assurance of service delivery					
Service specification	Identifying transaction and production cost characteristics of services (section 3.2)	Defining services as technical inputs and/or as business output (section 4.1, paragraphs 3-4 and section 4.2, paragraph 4)	Allocation of authority for significant decisions (paragraphs 1 and 4)		
Service level requirements specification	Costs of contracting (section 3.1, paragraph 3 and section 3.2, paragraph 3)	Input versus output monitoring (paragraph 4, section 4.2 paragraphs 2-5)		Measurement cost Sharing and value of information (paragraphs 5-6)	
Service level measurement and verification	Opportunistic manipulation of information (section 3.1, paragraphs 3 and 4)	Adverse selection and moral hazard Separation of decision management from decision control (paragraph 4, section 4.2 paragraphs 2-5)		Measurement cost Knowledge specialisation Information asymmetry (paragraph 3)	
Incentives for service level attainment					
Implicit incentive patterns	Incentive power of governance structures (section 3.1, paragraph 4, section 3.2, paragraph 5, section 3.4, paragraphs 9, 11-12)	The structure of compensation (paragraph 4, section 4.1, paragraphs 3 and 7, section 4.2, paragraphs 2-5)	Incentives accruing to asset ownership Incomplete contracting (paragraphs 3, 4-6, 10)	Information asymmetry (paragraphs 2, 4 and 7)	Reputation effects and corporate culture (paragraphs 8, 11)
Contractual reward/penalty devices		Incentive intensity, revelation of information, commitment and risk sharing (section 4.1, paragraphs 3, 4, 6, 7, 9, 11)			
Other contractual and organisational safeguards	Adaptation mechanisms (section 3.1, paragraph 4, section 3.4, paragraph 13)			Information sharing (paragraph 4)	Cultural match (paragraph 2)
Organisation and management of service provision					
Roles and responsibilities	Managerial adaptation (section 3.3, paragraph 7)				
Coordination and communication mechanisms		Influence and bargaining costs (section 4.2, paragraph 8)			Cultural match (paragraph 2)
Locus and scope of decision making	Adaptation mechanisms (section 3.3, paragraph 4 and sextion 3.4, paragraphs 8-9)	Separation of decision management from decision control (section 4.2, paragraph 4)	Relative importance of investments (paragraph 4)	Measurement cost Value of information (paragraphs 5-6)	

Table 5.1
The three layers of governance and the corresponding insights from the theories reviewed in chapter 2.

4.1. Assurance of service delivery

The purpose of this layer is to provide a specification of the boundaries of the services and their interdependencies, a characterisation of their economic attributes and the necessary performance measures and targets. In particular, the purpose here is to understand the nature of services in order to complete the groundwork needed for designing governance structures and mechanisms. This is achieved first by delineating the nature and scope of each service, second by specifying the required performance targets for each service and third by establishing ways of measuring and verifying realised performance.

4.1.1. Service specification

Proper service specification ensures that customers/users get what they need and ask for, achieve the desired level and quality of services, manage service provision and get the service provider to commit to specific deliverables.

IT services are not necessarily standard and uniform across companies, nor even across business units within the same company. What is important in any case is to specify each service in a manner relevant to user strategies, perceptions and requirements. Service specification can include either *what* must be delivered or/and *how* it must be provided. Defining what a service is to provide entails specifying the aims and the scope of serving the respective business functions, operations or processes. Apart from the aim and scope of a service, a company may wish to specify elements such as particular equipment or software, site-specific information and other detailed conditions, which relate more to how the service is provided rather than to what it must deliver. For example, network operations can be specified as the provision of a certain level of network availability, reliability, speed and security (what). Alternatively or simultaneously, it can be specified that a given number of people with particular skills will operate the pre-specified

equipment in a particular way (how). The distinction is between contracting for input or output[4].

There are four factors that must be considered when deciding between the two specification strategies. First, the determination of the material conditions of service provision (how), to great lengths, increases the control as well as the accountability of the client organisation over service delivery. This was explicitly recognised my management at BP. It may also inhibit adjustments and improvements over time as these would require contract renegotiation whereas a contract based on expected business services (what) is independent of the technological conditions of providing the service, which may change over time (see footnote 7). Second, if a company specifies only the aim, scope and success factors of the service (what), it delegates all responsibility to the service provider but also gives away more control of the service. Third, the risk associated with such decisions is usually borne by the party who makes the decision (see footnote 4). Finally, the feasibility and efficacy of performance measurement may constrain what can and cannot be specified[5].

---

[4] The basic informational properties of contracting for input and/or output are discussed as early as in the third paragraph of chapter 2. An equivalent argument is constructed in terms of the uncertainty involved in the provision of given services and the risk attitudes of supplier and customer (see chapter 2, section 4.1, paragraphs 3-4). The ability to ascertain whether a bad outcome is due to bad luck or little effort (when contracting is based on output performance) affects the efficacy of reputation mechanisms (see chapter 2, section 7, paragraph 9). With respect to positive agency theory (chapter 2, section 4.2, paragraph 4), contracting for inputs (how) relegates responsibility for decision making to the customer and responsibility for implementation to the supplier. Conversely, contracting for output (what) relegates responsibility for both decision making and implementation to the supplier.

[5] This is the main issue discussed in chapter 2, section 6, paragraph 1. Measurement problems are one of the causes of the incompleteness of contracting (chapter 2, section 3.1, paragraph 2, also section 5, end of paragraph 1) and a constituent part of secondary uncertainty (chapter 2, section 3.2, paragraph 3). The cost or efficacy of monitoring input or measuring output may also determine to some extent whether the contract is based on input (how) or output (what) and how risk and

BP recognised the desirability of emphasising the business aspect (what) of services but found it unavoidable in many instances to specify technical resources (how). See, for example, the opinion of BP America's VC of Operations and the initial intentions of BPX in chapter 4, section 4, paragraphs 5 and 6. BP America's VC of Operations emphasised the need to identify the necessary skills and the amount of involvement required (i.e. inputs) for the management of each service, whether it is provided internally or it is outsourced. As mentioned later in the same section, the top-level vision for IT in BPX emphasised the 'demand side' as opposed to the 'supply side'.

Nonetheless, they were aware of the implications mentioned above. BP call-off contracts variably specify inputs and outputs for each service or system. Input specifications include, among others, descriptions of the service (which may include specific software and hardware), interfaces with other systems, specific responsibilities and level of effort in person-hours for each task or skill-set involved. In addition, the charging and billing section of the contract often specifies specific budgets for particular activities (an additional measure to control inputs in cost-plus contracts). Output specifications include the scope of the service, who the users are, business criticality of various components of the service and service level targets. The two kinds of specifications are usually not separated in the contract. See chapter 4, section 7, paragraphs 4 and 10.

In identifying individual services, it is also important to take into account their inter-dependencies and complementarities[6]. There are many examples of service

---

residual claims are allocated (see chapter 2, section 4.1, paragraph 4 and related comments in footnote 4).

[6] Inter-relationships between tasks have repercussions for incentives and therefore for the allocation of tasks to agents. This issue is explained in chapter 2, section 4.1, paragraph 9. Inter-relationships between IT systems and services may give rise to the condition of asset complementarity, in which case residual rights of control for both assets should be allocated to the same party. See chapter 2,

inter-dependence, especially between end user applications and infrastructure systems. Database availability for example depends on both the availability of the server or host as well as on the availability of the network. Help desks are related to most other services, e-mail depends on LAN and WAN support and so on. Of course, all IT services are ultimately related in some way. However, an understanding of all inter-relationships is useful in pinpointing those that matter.

The contracts at BP did not explicitly address issues of inter-relationships between IT systems and services, although wherever applicable the interfaces of a given system or application were clearly stated. However, in discussion with managers at BP Oil it became evident that this was an issue of concern to them. Specifically, the issue of identifying the fault and the party responsible when a problem arose was a significant burden to them and a strong argument for total outsourcing. This highlights the systemic nature of IT service provision (see footnote 6).

Time is another parameter of IT service specifications. Services may change over time due to at least three influences. First, user requirements and technologies change regularly. Second, the evolving conditions of the IT supply industry also determine to some extent the shape of IT services. Third, service contracts are being evaluated and reviewed. For instance, if poor service quality and high costs prove to be due to inappropriate service specification, the latter would then have to be revised. Therefore, service specifications need to be reviewed on a regular basis and provisions for this may have to be made both in the contract and in terms of appropriate organisation and management processes to carry out the reviews (see

---

section 5, paragraph 4. Complementary IT systems and services effectively give rise to a situation of team organisation where the inputs ('team members') are a combination of different kinds of equipment and human skills. The joint output is more than the sum of its parts. Identifying such relationships in conjunction with the cost and efficacy of measuring the marginal contribution of each input has implications for the feasibility of different forms of organisation. See chapter 2, section 6, paragraphs 2 and 3.

section 4.3 further below)[7]. Provisions for changing service specifications over time depend on the broader governance context.

BP management found that moving from legacy mainframe systems to a client-server architecture is an important shift although it may not be wholly visible to the end-user. This shift also changes the economics of the contract if, for example, the outsourcer is responsible only for the servers and not for the desktops (clients). Running a data centre generates significant value to the supplier, whereas maintaining servers is relatively small business since value is now generated at the desktops. This is another case where service interdependence proves critical.

BP agreed with its suppliers an initial period of testing and fine-tuning performance measures, targets and incentive mechanisms before enforcing rewards and penalties or additional charges (see chapter 4, section 7.1, paragraph 5). This can be interpreted as a response to the need for *ex post* efficient adaptation while avoiding adverse incentive effects (see footnote 7). BP contracts also instituted formal procedures for handling Service Amendment Requests (SARs) from users. Moreover, for each system/service, the contracts prescribe particular change

---

[7] Contractual provisions create a self-enforcing range wherein contingencies are dealt with efficiently and effectively. Contingencies outside the self-enforcing range create opportunities for haggling and hold-ups. Additional management processes and decision making authorities need to be in place in order to deal with such situations and to revise contractual provisions. See chapter 2, section 1, paragraph 13. Adaptability to changing or unexpected circumstances is the main advantage of relationality in contracting (chapter 2, section 2.1) and of the legal regime governing internal organisation (chapter 2, section 2.2). The need to forego contractual promises when unexpected or changing circumstances require so, is the essence of the condition of imperfect commitment which permeates all theories reviewed in chapter 2 except mathematical principal agent theory (see chapter 2, section 4, paragraphs 1 and 4 and section 4.1, paragraph 2). However, there are merits to persistent, even if inefficient, commitment as uncontrolled adaptation may cause greater inefficiency due to adverse incentive effects. See the discussion on the ratchet effect in chapter 2, section 4.2, paragraph 6 and the economic theory of culture in chapter 2, section 7, paragraph 4..

control procedures for handling SARs specific to the given service. In most cases, the contracts require the supplier to introduce revised change control procedures for each system/service. See chapter 4, section 7, paragraphs 8-9.

Service specification is largely subjective. It is a matter of how the users, the IT function and the general management want to perceive the purpose of each service. For example, help desks and voice telecommunications are closer and more visible to the users. Database support and back-ups on the other hand are more technically oriented services although the ultimate goal is always to satisfy end customer needs. Many parties within an organisation are likely to have a stake in the specification of each service and their interests are likely to diverge or even conflict. Service specifications may be stretched to cover all different perceptions. However, for purposes of contract clarity, efficiency and enforceability, it may be preferable to agree on a narrow and well-demarcated specification. Otherwise transaction costs will rise significantly: specifying, writing, monitoring and enforcing an explicit or implicit contract for broadly specified services will be very costly. In practice, it will be difficult for the supplier to 'serve many masters' and for the customer(s) to assess performance. This is not to say, of course, that important facets of a service should not be included in its specification.

BP call-off contracts typically specify who the users are and who the relevant contacts from the supplier's side are for each system/service. This and other similar measures to communicate the new structures were taken to instil confidence to BP users as to who their IT suppliers would be. User considerations informed output specifications (what) whereas technical and financial considerations informed input specifications (how). However, a supplier site manager raised a concern that coordination structures created too many interfaces between BP and the supplier. The latter effectively had to serve too many masters with often conflicting requests at the same site. See chapter 4, section 7.2, last paragraph. This also points to the fact that contractual relationships do not rely exclusively on static documents (i.e. the contract) but are continuously adaptive processes through which users and

providers constantly 'negotiate' their exchange. This is one important aspect of relational contracting (see chapter 2, section 2.1, paragraph 4).

Service specification should, finally, include several characteristics including the degree and desirability of existing asset specificity, the level of demand and its variations over time (as an indication of resource utilisation), the robustness and maturity of the technology (as one indication of uncertainty), the minimum efficient size of operations (potential scale economies) and resources shared with other services (potential scope economies)[8]. Clearly identifying and estimating these characteristics is a major task and it may even be impossible or too costly in practice. Practical difficulties aside, it is important to generate some relevant indicators, either absolute (concerning individual services) or relative (comparative service ranking).

BP did not consider these issues explicitly for individual systems and services. The general strategic mandate at BPX was, implicitly, that these parameters were such that total outsourcing was the preferred option. In discussions with managers, however, it became apparent that opinions diverged. BP Oil certainly preferred a more incremental approach, in recognition, again implicitly, of the varying constraints posed by these parameters. Having said that, they extensively referred to the notion of the 'maturity' of a technology or service, to signify the uncertainty associated with it, the extent to which it is become a commodity service or not and/or the extent to which they felt a fairly discrete (hands-off) contract for that service would be efficacious. This directly related to their perception of the ease of specifying performance measures and of their effectiveness. See chapter 4, section 7.1, paragraph 15.

---

[8] These are the basic parameters determining production and transaction costs. See also chapter 2, section 3.2.

At the first presentation I made to a group of senior IT and business managers from BPX and BPO, I tried to explain, among other things, the concept of asset specificity in fairly general and intuitive terms. Whereas most of them never quite mastered the pronunciation of the term, once they grasped the concept, they kept using it regularly even after I had completed the case study. The acceptance of specificity in regular discussions about organising IT (as well as other things such as oil rigs and convenience stores at petrol stations) testifies to the power of the concept, despite accusations of it being elusive and tautological (see chapter 2, section 8.1, paragraph 15).

To summarise, the elements of governance highlighted in this sub-layer include the specification of services in terms of input or output, the interdependence and complementarity among IT services, the need for adaptations on contractual requirements and the multifaceted characteristics of services. The strategic direction of BPX was to concentrate on demand management, rather than supply management. However, managers throughout the Group were sceptical as to the possibility of ignoring technical issues (input specifications). The structure of contracts shows that this was not possible.

The outsourcing exercise at BP also called for seamless service provision, emphasising the need to avoid the burden of coordinating different services and different suppliers. They took several measures in this directions but were not able to satisfy their aim completely. Of particular concern to BP management was the need for flexibility in IT contracts. They had already experienced difficulties in getting out of inflexible contracts when the technology improved or the business changed. BP managers were also implicitly aware of the different characteristics of IT systems and services that facilitate the implementation of different kinds of contracts. However, their only operationalisation was a poorly defined notion of maturity. All these issues are examined above in a systematic manner by analysing the contractual evidence at BP and by drawing on existing relevant theories.

4.1.2. Service level requirements specification

While service specifications circumscribe the business and/or technical purposes served, service level requirements (SLR), the quantity and quality to be delivered, constitute the operationalisation of these purposes. From a contractual perspective, service level requirements specify the objectives that the provider (internal or external) is contractually (formally or informally) bound to meet. They therefore constitute the central component of the contract and as such have an impact on the costs of contracting and consequently on the form of governance.

SLRs comprise two, often indistinguishable, components: the metric or measure and the respective target value. For example, a metric for PC support in one of BP's contracts is "*response to severe problems*" and the target "*Only one problem can take longer than 30 minutes to fix or refer to the supplier in any calendar month if there are 20 calls or less, or two if there are more than 20 calls*". See chapter 4, section 7.1, paragraph 2.

SLR metrics and targets can be described in either business (e.g. percentage of successful video conferences) or technical (e.g. database hours or person-hours) terms, or both. The distinction is parallel to the aforementioned trade-off between specifying *what* a service does and *how* it does it and therefore SLRs partly depend on, as well as determine service specifications. More specifically, business oriented SLRs ensure that end-users receive the required service by passing the full responsibility of accomplishing it to the provider. The pitfall is that such metrics are susceptible to changes in the business and the user requirements. With technically oriented SLRs the customer/user needs to maintain adequate technical knowledge in order to set and monitor the requirement. The customer/user is also held responsible for the particular choice. On the positive side, technically oriented requirements secure the availability of certain resources and/or procedures in the event of contingencies or business changes. Moreover, if there are several technical

alternatives for a given service, technical requirements are a means of cost control. See footnote 4 above.

Although BP wanted to separate IT provision as a business service from technology management (as mentioned in the previous section), they faced serious difficulties in the process of defining service level requirements. Moreover, they explicitly recognised the above mentioned implications of setting business oriented or technology oriented SLRs. See chapter 4, section 7, paragraph 10 and section 8, paragraph 7.

The efficacy of service level requirements can be assessed on the basis of two criteria, namely, completeness and precision. *Completeness* characterises the extent to which a service level requirement covers all aspects of intended performance. Service level requirements for each service need to be as complete as possible because even minor omissions may lead to very low overall service performance. On this matter, managers at BP often quoted the case of voice services as example. Specifying that 90% of all incoming calls must be answered within 5 seconds is not enough if there is no provision for the remaining 10%, which may in practice never be answered. *Precision* describes the extent to which a given SLR is an accurate and unambiguous measure of actual performance. For instance, user satisfaction surveys for measuring vendor performance in desktop support, are highly imprecise because user satisfaction depends on many factors outside the vendor's control and because the process is subject to various sources of bias. In contrast, network downtime can be measured with perfect precision[9].

---

[9] The role of completeness and precision in SLRs is to control uncertainty resulting from measurement problems (chapter 2, section 3.2, paragraph 3), to alleviate information asymmetry and hence, to contain opportunism (chapter 2, section 3.1, paragraph 3 and section 4, paragraph 4) and to support the operation of incentives by focussing the attention of the supplier to the right aspects of performance.

Discussions with managers at BP highlighted the limitations and gaps in service level requirements. They considered their service level agreements to be good and the grief point system to be important. However, they did not hesitate to emphasise the fact that gaps remained in the specification of metrics and in the grief point system and that SEMA was able to exploit such gaps by sticking to the letter of the contract. Near the end of the case study, BP started an effort to introduce broader (i.e. more complete) metrics addressing the overall quality of service (as opposed to measuring narrow performance indicators) to remedy the problem. See chapter 4, section 7.1, paragraphs 10 and 17 and section 8, paragraph 8. These observations lead, first, to the notions of completeness and precision as the defining qualities of service level requirements and, second, to the notion of the lifecycle of service level requirements explained below.

Further repercussions of failing to specify complete and precise service level requirements accrue from service inter-dependence. Inasmuch as service level requirements are the basis of incentives that affect the provider's performance, poor performance due to poorly specified service level requirements on one service may have knock-on effects on other related services. For example, no matter how well the metrics of all services are defined, if the metrics for network operations are incomplete, resulting in low performance, the whole IT service provision will be affected (see footnote 6 above). Management at BP were particularly aware of this and made explicit contractual provisions to allocate responsibility for faults between different vendors (see chapter 4, section 8, paragraph 10).

In general, IT services have various facets (functions), for each of which, a set of appropriate service level requirements can be devised. For example, facets of help desk services in one of BP's contracts include solving problems quickly, keeping the customer informed of the progress made, referring problems to the groups responsible for complementary services, etc. For each of these facets, appropriate SLRs must be specified as completely and as precisely as possible. Completeness is

therefore achieved by taking account of all facets of a service and by avoiding loopholes in specifying service level targets.

Maximising the completeness and precision of service level requirements may not be cost effective in practice. Technical complications and measurement difficulties of various kinds may inhibit the achievement of perfect completeness and precision. User satisfaction surveys are an obviously desirable performance measurement instrument but it is difficult to base incentive payments on it due to the biases mentioned above. Air BP actually implemented such an incentive but it was not considered particularly effective. Other parts of BP did not attempt it (see chapter 4, section 7.1, paragraphs 11 and 15). In addition, the process of administering such surveys is too costly to be carried out on a frequent basis, since it requires the time and effort of many individuals. The number of successful videoconferences is a potentially easy to measure (and possibly to verify) metric. However, 'success' must be defined in the contract. These examples demonstrate that the development of service level requirements involves costs of various kinds. The different kinds of costs involved can be understood on the basis of a lifecycle of SLR development. This lifecycle emerged during the case study as a useful tool for debate on devising and assessing SLRs (see ch4, section 2, paragraph 13).

4.1.2.1. Lifecycle of service level requirements

It is useful to conceptualise the development of service level requirements as a lifecycle (Figure 5.5). Each stage poses varying difficulties. For example, user surveys are easily identified as being useful, but are more difficult to specify in detail, to write a contract for and to implement. In contrast, system response times are relatively more easily developed throughout the lifecycle.

The first step (identification) is the selection of desirable and feasible metrics for each service. This process involves either choosing among commonly used metrics or designing novel ones, or both. Some desirability criteria are relevance (is it

informative of the performance of the service or of its contribution to the business?) and necessity/usefulness (does it measure an important or critical aspect of the service? - does it really have to be measured?).

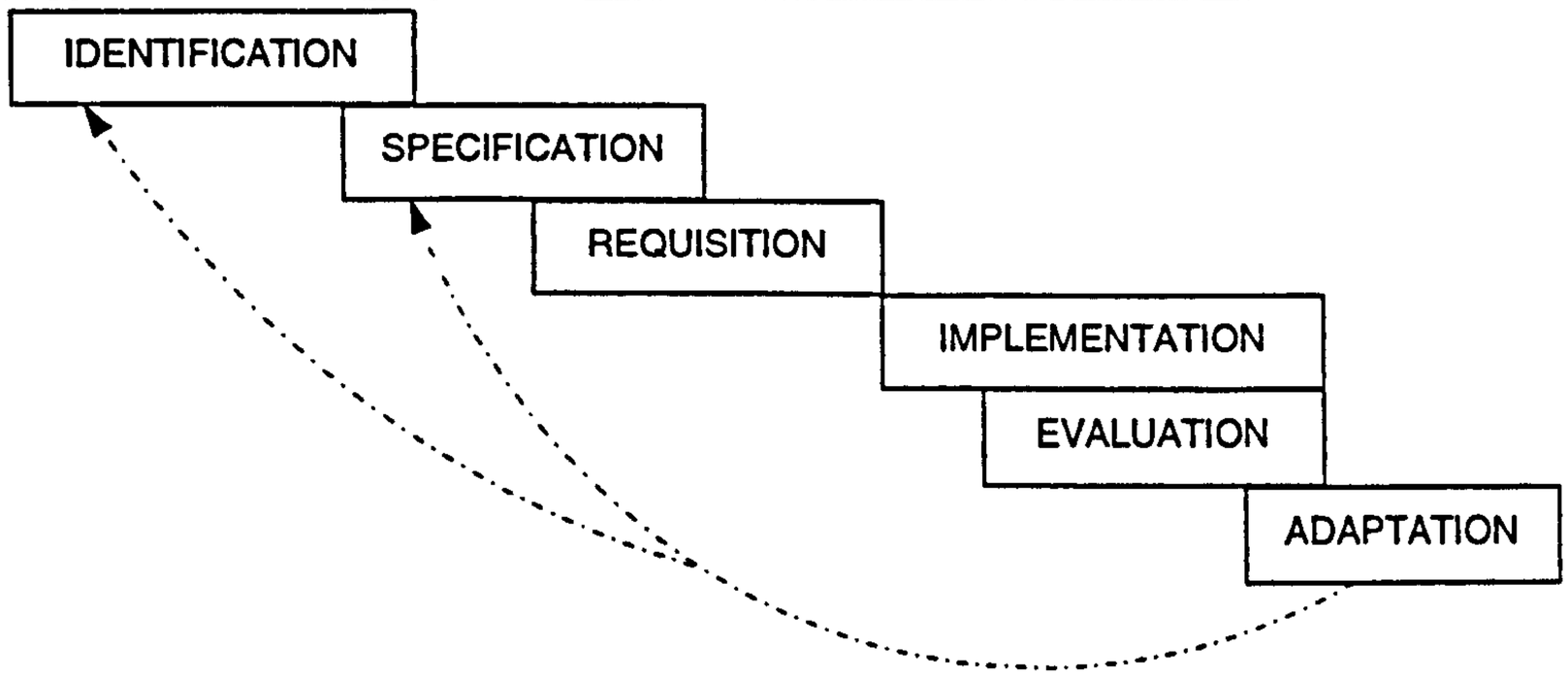


Figure 5.5

Lifecycle of Service Level Requirements

The need for a service level requirement is first identified and defined (Identification). The associated performance targets are then specified (Specification). SLRs are then negotiated, included in the contract (Requisition) and implemented during contract execution (Implementation). At the same time they are also evaluated, whether explicitly or implicitly, so that weaknesses and gaps may surface (Evaluation). Appropriate measures are then taken to adapt existing SLR metrics and targets or to introduce new ones (Adaptation). This is an iterative process throughout the life of a service.

Potential service level requirements can also be judged against their feasibility and cost of development through the rest of the lifecycle (ease of specification, ease of requisition and ease of implementation).

In the case of desktop support in one of BP's contracts, for example, one metric related to service quality is the availability of adequate skills to advise users. This metric is certainly desirable (relevant and necessary) but it might be difficult to specify in advance exactly what these skills have to be and how their availability should be ensured, given the prospect of changing desktop applications or systems over time. BP implemented this metric, being aware of the fact that they would have to revise it regularly.

Having selected the metrics that are going to be used for a particular service, their target values have to be determined (specification stage). This may involve benchmarking existing service levels and specifying the required level and/or a plan for improving it. BP business units employed various tactics in specifying performance targets. See chapter 4, section 6, paragraph 3 and section 7, paragraph 8. The means of measuring and verifying the set target should also be specified. In the case of help desks, for example, response times are often specified contingent on ranges of demand levels. Unquantifiable metrics are verified by administering user surveys, whereas more tangible ones are directly measured. BP (see chapter 4, section 7.1, paragraphs 4 to 6) employed various combinations of such approaches, with a view to carrying out effective performance monitoring while minimising the costs of doing so.

Setting targets for existing services is relatively easy as compared to new, especially innovative, services (e.g. if there is no or little relevant experience in the industry). This process is even more problematic with client-specific services, since by definition there is little or no similar experience in the market. In the absence of comparable measures for specifying service level targets, more emphasis might be placed on serving the business requirements.

The next stage (requisition) corresponds to the actual inclusion of service level requirements in the contract. The contract may be formal or informal, explicit or implicit. In any case, the company and the service provider (whether this is an employee or another company) will probably have to negotiate over the terms of the introduction and enforcement of each metric. This is also the stage when contractual technicalities are resolved.

A service provider would not be willing to contract for service level requirements that specify unachievable targets and would attempt to negotiate. Moreover, a service provider would not accept full responsibility for less tangible and objectively verifiable metrics such as user satisfaction, which are influenced by a

variety of factors, including influence activities and opportunistic manipulation of information.

The first three stages (identification, specification and requisition) are often indistinguishable in practice, as they are resolved simultaneously through negotiations with the provider. Even so, this conceptual distinction helps to highlight important aspects of the desirability and costs of SLR development.

During contract execution (implementation) IT services are actually being delivered and performance is realised, measured and verified. Service level requirements are implemented and simultaneously evaluated (fifth stage). While the service is being delivered and the metrics implemented, various pitfalls in the original specifications and targets may appear. Quite often unanticipated side-effects distort initial intentions.

For this reason, problem monitoring and reporting procedures are established that extent beyond the measurement of the set performance metrics. The rationale is to capture aspects of service quality which were not anticipated in advance and thus ignored in the contract. This is often formally agreed to take place over a fixed initial period, during which all arrangements are being tested and fine-tuned. Agreeing on such a fixed period helps economise on excessive monitoring and bargaining costs while providing an opportunity for desirable adaptations. BP agreed to an initial period for evaluating and adapting service level requirements (chapter 4, section 6, paragraph 4 and section 7.1, paragraph 5). Furthermore, they agreed to the reporting of additional information, alongside the service metrics, in order to qualify the nature of the performance shortcomings observed in the primary metrics (chapter 4, section 7.1, paragraph 3).

According to the experience gained from the implementation and evaluation of a set of service level requirements, the company may have to revise some or all of those metrics and may introduce new ones, if such adaptations are desirable (i.e.

they do not involve costly renegotiation or adverse incentive effects) and permissible (by the contract and the law).

The process of the SLR lifecycle is iterative and ongoing throughout the life of a service. The governance, the contracts and the supplier associated with a given service may change but the SLRs for this service can survive such transitions. Service level requirements can, therefore, be seen as a property of the service rather than just a feature introduced by the contract.

Although service level requirements need to be as efficacious as possible, the difficulties (costs) of developing them through their lifecycle should also be taken into account. In general, one should continue improving the effectiveness of service level requirements for each service until the incremental effort for further improvement at all lifecycle stages exceeds the respective increase in efficacy (completeness and precision).

Of course, in practice it is almost impossible to calculate these marginal costs and benefits, although it should be clear when difficulties cannot be surmounted and when SLRs are inadequate. It may be necessary to reconsider service specifications (define services in a different way), or devise alternative governance mechanisms (such as direct supervision) that ensure effective service delivery, if either the effectiveness of service level requirements is judged unsatisfactory, or the cost (difficulty) of developing the service level requirements through their lifecycle is too high (i.e. the metrics are not easily specified, negotiated, measured and verified).

Obviously, the costs of administering service level requirements must be kept to a minimum because they create no value (chapter 2, section 1, paragraph 8). Their purpose is to enforce the delivery of the required service levels, or in other words to ensure that the parties keep their promise regarding service levels and costs. If parties could rely on altruism and good faith, they would probably avoid using

service level requirements. This was explicitly stated by management at BP, with reference to the usefulness of performance reports *vis-à-vis* the effort consumed in producing and reviewing them (chapter 4, section 8, paragraph 12). BP employed a variety of measurement and reporting methods in order to minimise the respective costs (chapter 4, section 7.1, paragraphs 4 and 6). Because of these costs, moreover, some performance requirements were not assigned rigorous targets but only qualitative expectations (chapter 4, section 7.1, paragraph 8). Decision making authority and corporate culture may be able to relieve, to a certain extent, these problems (see section 4.3 below and chapter 2, section 7, paragraph 2).

Having said that, companies can also generate value from this effort by taking advantage of the information generated. This can be achieved by designing metrics that are potentially rich in informativeness and by setting up appropriate organisational structures and processes that will exploit this information for improvement, innovation and, generally, better decision making. If, for example, monitoring information is also used to improve a service or reorganise its provision, then it does have productive value. One should try to minimise the development cost of SLRs while maximising the potential value of the information collected with them. As mentioned above, this is what BP management had in mind when they debated the usefulness of performance reports. The productive utilisation of this information will in turn require appropriate organisational structures, which are discussed in section 4.3.3.

To summarise, this section examines the use of service level requirements in BP contracts, under the light of relevant theories. In particular, four main issues are highlighted. First, the distinction between the metric and the target was evident in all BP contracts. Second, completeness and precision as key quality criteria of SLRs. Originally, a long list of desirable qualities for SLRs was being contemplated with BP management during the case study. After reviewing the concerns raised about the efficacy of the SLRs already in place, the list was condensed to completeness and precision. Third, discussions with BP managers on the difficulties and benefits of

developing and administering SLRs lead to the notion of the lifecycle. It was seen as a useful tool for contemplating the 'costs' and 'benefits' of SLRs. Fourth, a brief reference to the measurement of SLRs and the use of performance information was prompted by a growing interest in BP to derive value from such information. In addition, this section makes references to the issue of defining services in technical or business terms and to the issue of service interdependence. Both these issues featured in the previous section.

4.1.3. Service level measurement and verification

The process and frequency of service level reporting are an important part of SLR specification. Keeping track of service delivery entails a number of costs and has incentive implications. Therefore, it matters who will undertake it and how[10]. In addition, ensuring that reports are truthful may prove to be a problem if customers and service providers have an interest in misrepresenting information. In order to ensure the delivery of the required service levels and to rely on metrics for incentive and control purposes, appropriate measurement and verification mechanisms need to be instituted in advance.

Two types of information hiding or misrepresentation are possible. First, suppliers are often in a position to overstate their costs. Second, they may distort information about the difficulties encountered or the capabilities and costs of technology. Either problem may arise during supplier selection (adverse selection) or after the contract is in place (moral hazard)[11]. Incentives, monitoring, auditing,

---

[10] Monitoring, as part of decision control has incentive implications for positive agency theory (chapter 2, section 4.2, paragraphs 2-4). In order to counter agency costs, the bearer of residual claims should also exercise decision control (ratification and monitoring). In IT outsourcing contracts matters are more complicated than the stylised case presented by the theory. See chapter 2, section 4.2, paragraph 5.

[11] Moral hazard and adverse selection are examined in chapter 2, section 3.1, paragraph 3, and in chapter 2, section 4, paragraph 4.

benchmarking and third party experts are means that can be used individually or in combination to address these problems. Their cost and efficacy depend on the circumstances. Monitoring, measurement and verification are the subject of this section. Incentives and other safeguards are discussed in section 4.2.

In general, there are obvious reasons why both customers and service providers would have an interest in concealing or misrepresenting the truth about the delivery of service levels. Customers would be able to ask for more (or avoid payment or claim penalties) if they understated the service levels that they actually received. Service providers would be able to accumulate excess profits by overstating their achievement (hiding poor performance) or the cost of service. This is possible due to conditions of information asymmetry and opportunism (information impactedness)[12]. Such opportunistic tactics are not often employed in extreme forms which would lead to severe conflict and even contract termination, but rather in small incremental steps thus expropriating any information rents that remain beyond the efficiency of incentives and monitoring (residual loss - see chapter 2, section 4.2, paragraph 1). They also appear in different patterns, depending on the duration of the relationship and the associated uncertainty. As noted by Muris (1982), opportunistic behaviour is subtle because it is easily disguised as legitimate and its opportunistic nature can be discovered and verified only at a cost.

If the service levels delivered are symmetrically observed by both parties, there is little space for opportunistic manipulation of information. This may be the case with service levels that have an immediate impact on business processes, or on everyday work, such as desktop support, database availability or network downtime. The performance of other services may not be as apparent to the

---

[12] When the parties do not have access to the same information (on costs, performance or other conditions) or when objective means of ascertaining that information are absent or too costly (such as a third party expert), opportunism leads to misrepresentation and haggling. This is the concept of information impactedness (chapter 2, section 3.1, paragraph 4).

customer (e.g. systems development, project management and security), or may become apparent after a significant time lag (e.g. when it is too late to salvage a project or when a security breach causes obvious damage). However, that such information is symmetrically observable does not necessarily mean that opportunistic tactics are completely avoided. Nonetheless, the availability of this information is necessary in order to take further action (e.g. to revise incentives or to introduce penalties).

Of course the cost and effort of service provision is known only to the provider and, when the latter is external, various mechanisms such as audit rights and benchmarking can be instituted to alleviate some of the effects of information asymmetry. BP management were not particularly concerned with possible misrepresentation of information. Despite the lack of regular mechanisms for verifying the suppliers' performance reports, they were confident with their ability to detect and address misrepresentations. The framework agreement specifies an 'open-book' policy and gives BP the rights to audit and benchmark the activities of their suppliers (chapter 4, section 7, paragraph 1 and section 7.1, paragraph 6).

The costs of measurement and reporting depend on the nature of the metrics used and have to be kept to a minimum. Service clustering may generate scale economies in monitoring and measuring service levels thus affecting the total cost of governance. There are three main types of metrics in increasing order of measurement cost. (i) automated metrics measured by software or hardware modules built into the various systems, (ii) standardised activity logs kept by individuals and (iii) customer/user surveys or review meetings which may or may not be standardised and require the involvement of a larger number of people. BP's contract employ combinations of all three kinds of measures (chapter 4, section 7.1, paragraphs 4, 11 and 12).

Service level measurement and reporting can be carried out by the user/customer, the service provider or a third party (e.g. auditor). The decision will depend on the

informational properties of the service cluster itself as well as on the procurement strategy being contemplated. The aim is to minimise the cost of measurement as well as the agency costs of moral hazard and adverse selection. If customers cannot observe service levels directly, they must be able either to induce providers to report truthfully by offering a contract that rewards honesty or penalises cheating by more than the latter would pay, or to rely on direct monitoring or on third party intervention (auditing, arbitration or the courts) for resolving disagreements[13]. For example, in an outsourcing contract, the supplier might undertake service level reporting because they are in a better position to observe their own performance (and thus economise on measurement costs). Otherwise the customer would have to expend resources for monitoring. Periodic verification, audit rights and incentives are then needed to induce truthful reporting. This is the approach followed by BP (chapter 4, section 7.1, paragraph 6).

The costs of monitoring can be contained while maintaining or even improving incentives by reducing the frequency of measurement or the amount of information collected and by simultaneously setting a large fixed reward or penalty (Cowen and Glazer 1996). For example, measurement at random intervals combined with severe penalties for underperformance create a powerful incentive that can be employed for non-critical metrics and/or when measurement is too costly. For example, BP has explicitly classified certain metrics as 'mystery shopper' (chapter 4, section 7.1, paragraph 4). These metrics are measured at random intervals. For example, a 'mystery shopper' metric for help desk services is "*contact a sample of users to confirm problem handled satisfactorily*". Another example, for e-mail administration and support, is "*requests for all mail IDs or disks to be*

---

[13] When SLRs are not a complete and precise representation of reality or when information about the conditions affecting the supplier's performance is not readily available, direct monitoring of the supplier's effort is needed (e.g. auditing). If there is a competitive market for the given service(s), then market benchmarking may be a sufficient source of information. These measurement problems are examined in chapter 2, section 6, paragraph 3. The incentive mechanisms inducing truthful reporting are examined in chapter 2, section 4.2, paragraph 5.

actioned within two working days". However, no penalties are attached to them. Clearly, the frequency of measurement and reporting will depend on the characteristics of each SLR, including their cost of measurement and incentive consequences.

Monitoring and reporting of service delivery may include other information in addition to the specified service level targets (e.g. progress information, exception reports, persistent problems, etc.). Clearly, keeping track of more information entails added costs and its necessity should therefore be scrutinised. Additional information monitoring is justified in three instances. First, during the early stages of the contract, any aspects of service performance not captured by the set metrics will be highlighted and taken into consideration. Second, if the parties have little experience with a given service, if the supply relationship is still young and 'immature' and if the service is characterised by a high degree of uncertainty, there may be learning benefits to be had from the exchange of additional information. Finally, in cases of specific services where tacit organisational or other service-specific knowledge constitutes an important asset, the exchange of additional information serves the transfer of this knowledge. BP collects and reviews such additional information, mainly to address persistent problems and to revise service level requirements (chapter 4, section 7.1, paragraph 6 and section 8, paragraph 8). IT also appears to be the case that some BP managers felt uneasy relinquishing control to the supplier. This was the opinion of the SEMA Account Manager (chapter 4, section 7.2, paragraph 4).

To summarise, this section addresses the measurement and reporting of vendor performance information. Managers at BP often gave the impression that this was seen as a trivial, if not burdensome, activity. However, the contracts made extensive provisions for ensuring that all information can be collected and distributed to all interested parties. Such provisions include the supplier's obligation to collect, distribute and review performance reports and BP's right to audit and benchmark the vendor's activity. BP collects additional information in order to ensure the

correct interpretation of recorded performance, to detect persistent problems and to identify opportunities for improvement. For BP, the possibility of misrepresented performance information was not an important issue. All the procedures mentioned above were perceived to address this problem satisfactorily. Vendor opportunism was manifested in other ways, as discussed in the following sections.

4.2. Incentives for service level attainment

The need for incentives arises mainly because the IT service supplier and the customer typically have divergent interests. The purpose of incentives is therefore to align these divergent goals, or in other words, to reorient the efforts of the supplier (agent) towards the goals and interests of the customer (principal) who sets the incentives.

There are three main incentive instruments of interest to the design of IT governance. First, intrinsic incentives accruing to the organisational mode and the allocation of asset ownership. Second, formal incentives tied to various measures (e.g. service level or cost). Third, contractual obligations (e.g. 'open book' policies) and organisational safeguards (e.g. corporate culture).

The primary source of incentives is the basic compensation scheme. Commercial contracts are mainly distinguished into fixed-cost and cost-plus contracts. The former provide greatest incentive intensity by making the supplier the residual claimant. The latter generate no incentive but may still be desirable in the presence of significant uncertainty. Mixed mode compensation schemes are also not uncommon, particularly in public procurement and regulation. Such contracts may place an upper limit to the cost component and/or include an additional performance-based incentive[14]. In internal organisation, salaries may be

---

[14] Chapter 2, section 4.1 examines the choice between cost-based and performance-contingent compensation schemes. The vendor's limited liability (risk aversion) increases the cost-based part of

supplemented with various types of performance-based bonuses (e.g. share options). The implicit promise of career rewards (which, of course, vary from one firm to another) is the typical incentive of the employment relation.

Most contracts that BP entered into, define total compensation as a combination of a fixed element, a cost-based element and a performance contingent element. The relative weights of each element vary from one contract to another (chapter 4, section 7, paragraph 5 and section 7.1, paragraphs 15 and 20).

Incentives are generally designed to ensure that service levels and cost targets are met. One of the goals of the IT function is also to pursue continuous innovation both in the delivery of IT services and in the contribution of IT to the business. Incentives that motivate improvement and innovation are certainly desirable. Incentives for innovation pose particular problems due to the nature of innovative activity.

Innovation in IT can be of two kinds. First, general innovations, not specific to the users and second, innovation that generates value only within a specific user environment. Incentives for innovation of the first kind are more easily created in the case of external IT provision, if, for example, the supplier receives a share of any cost reductions across a number of client organisations. Streamlining of operations and standardisation of systems are actions often taken in this respect. However, replacement of a system specific to a customer for a standard alternative is likely to lead to loss of the value accruing to specificity and specialisation. The customer affected will have to assess this loss against the respective cost reductions. Such initiatives are subject to supplier opportunism, if the alternative technology is

---

compensation and thus reduces incentive intensity (paragraph 3), despite any fixed fee proportions and/or rewards and penalties in the compensation structure. Uncertainty also increases the cost-based part of the supplier's compensation, even under asymmetric information about the extent of uncertainty (paragraph 7). The structure of compensation also has implications for input or output contracting (see section 4.1.1, paragraph 4 above).

specific to the supplier. Furthermore, problems in valuing the costs and benefits of such changes may, in turn, lead to costly negotiations and haggling. This prospect may deter the supplier from taking such initiatives in the first place. Auditing, benchmarking, escalation processes, third party experts and alternative dispute resolution procedures are examples of remedies to valuation problems (chapter 2, section 3.4, paragraphs 13-14).

Incentives for user-specific innovation are more difficult to provide since, by definition, the user is the residual claimant. In the presence of unsurmountable valuation difficulties, the customer would have to commit to award any such innovations as additional business to the innovating supplier for a sufficient time period[15]. A monopolistic situation emerges which may not be comparatively efficient. Moreover, the incumbent supplier is likely to innovate only within the bounds of his extant capabilities. Better incentives for innovation beyond these limits can be provided to the users themselves internally[16]. Apart from the typical incentive problems faced by internal organisation in this respect, the acquisition and maintenance of the skills and expertise needed (which might be a reason for outsourcing in the first place) are an added difficulty.

---

[15] Such a commitment is a type of hostage, as explained in chapter 2, section 3.4, paragraphs 11-12. In order to proceed with a client-specific investment, the supplier needs an assurance that he will not be left stranded. Assurances (hostages) can take various forms but they often create undesirable incentive side-effects (windows for opportunism). Besides, such commitments may not be efficient later in time, if the technologies or the markets change.

[16] This is the message of the theory of property rights in chapter 2, section 5, paragraphs 5-6. Ownership is the ultimate form of commitment and alleviates the incentive problems mentioned in footnote 15 above. Moreover, client-specific expertise is native in internal organisation, whereas an external supplier has to invest in acquiring and maintaining this knowledge (chapter 2, section 6, paragraph 7). However, internal organisation faces other incentive problems and it may not be worth acquiring the technical expertise of the supplier. Therefore, the choice between bilateral commitment and internal innovation depends on the particular circumstances.

Innovation and continuous service improvement has been an important issue for BP. One of the main concerns raised was that suppliers remained faithful to the letter of the contract and avoided introducing such innovations and improvements. BP wanted to introduce additional monetary incentives to this effect (chapter 4, section 7.1, paragraph 19). However, attempts by SEMA to standardise some systems and applications were resisted by BP as it was felt that significant user-specific value would be sacrificed (chapter 4, section 8, paragraph 15). Clearly, the incentives of SEMA to reduce costs by standardising are not perfectly aligned with BP's desire to increase value through specific improvements.

4.2.1. Implicit incentive patterns

A fundamental condition that determines the incentive characteristics of different organisational modes is the availability and sharing of information. In the case of internal IT provision, organisations can employ authority and fiat to observe all the information relevant to service delivery: the costs, the technology, the service levels (see chapter 2, section 2.2, paragraph 4). Despite the obligations imposed by the labour law, this is not to say that employees always behave in the best interest of their company and cannot withhold information. On the contrary, appropriate incentives are still required, whether formal (e.g. performance bonuses) or implicit (e.g. organisational culture), despite the fact that incentives within a hierarchy are necessarily less acute than those available in market exchanges[17].

Market transactions such as off-the-shelf hardware and software purchases, are characterised by a different incentive and information pattern. The incentives for high quality and low cost are primarily determined by the competition in the particular supply market. Perfectly competitive markets rely on the price system to

---

[17] Because of the collective nature of information in teams, team members have an incentive to shirk (chapter 2, section 6, paragraph 2).

convey all necessary information efficiently. However, such markets are rarely, if ever, observed in practice[18].

In supply relationships in the middle range between pure markets and hierarchies, the authority and discretion of each party are determined by the respective contract(s) and the law. Depending on how IT services are specified, it may be more or less costly to measure and verify the service levels delivered[19]. The effort and generally the costs of providers are known only to them, whereas the quality of services may or may not be observable by either.

For example, BP found that suppliers chose not to share information and best practices across sites, since they were competing for a larger share of business with BP. The additional quality incentives they were contemplating in early 1995 included criteria pertaining to the collaboration and information exchange between BP and supplier staff in formal and informal communications (chapter 4, section 7.1, paragraphs 17 and 18).

Reputation is one of the strongest implicit incentives in long term and repeated relationships. Opportunistic tactics sooner or later trigger a retaliatory response from the other party thus initiating a vicious circle of opportunistic moves. In this scenario, even if either party 'wins' at the end of the 'game', they both loose a potential, mutually beneficial trading relationship. In long term relationships, one or both parties will cheat only if the benefits from cheating are larger that the total benefits from the whole life of the relationship. This is generally unlikely and would only appear towards the end of the contract. This is not to say that long term

---

[18] Given the imperfections of markets, information acquisition is a widely accepted justification for vertical integration (chapter 2, section 1, paragraph 3). Private and asymmetrically available information creates adverse incentive effects, namely adverse selection and moral hazard (chapter 2, section 3.1, paragraph 4 and section 4, paragraph 4).

[19] The extent to which access to and gains from information can be contracted determines the governance structure (chapter 2, section 6, paragraph 4).

supply relationships are safe from opportunistic behaviour. Even if the two parties are not engaged in a long-term contract, the first-time looser will either attempt to retaliate in possible future exchange or avoid repeating trade altogether. Market reputation usually tends to be an effective deterrent of such intentions[20]. BP managers recognised that suppliers face strong incentives from the need to maintain good market reputation, independently of the rewards and penalties stipulated in the contract. Moreover, BP is a big client with significant purchasing power and the potential to offer more work to its suppliers over time. Having recognised that, they also admitted that suppliers behave opportunistically in incremental ways, without causing major contract breaches. See chapter 4, section 7.1, paragraphs 20 and 21.

Cases in which one of the parties (either the customer or the provider) has a very large comparative bargaining power and is in a position to negotiate and impose unfavourable conditions to the other, appear to be exceptions to the above discussion. There is a subtle difference however, in that in the latter case, the conditions are negotiated and voluntarily agreed upon, whereas in the first case, the parties attempt to take advantage of gaps in the contract after the agreement is concluded[21]. Small, incremental shirking attempts are more likely. Either party will

---

[20] The basic way in which reputation operates as a governance mechanism is explained in chapter 2, section 7, paragraph 8. Reputation is often based on personal relationships between supplier and customer staff but also extends beyond the bilateral relationship to the broader market or community (chapter 2, section 2.1, paragraph 5). The strength of reputation effects determines to some extent the cost (and efficacy) of hybrid governance structures (chapter 2, section 3.4, paragraph 9). Strong reputation effects make long term (or repeated) contractual relations more sustainable. Otherwise vertical integration is the preferred governance structure.

[21] This is the meaning of opportunism , as defined in terms of appropriable quasi rents and in terms of moral hazard. Appropriable quasi rents are a measure of the unredeployable specificity of an investment, which creates the possibility for a hold-up after it is committed (chapter 2, section 1, paragraphs 10-12). Opportunism as moral hazard also arises postcontractually as a result of asymmetric information. Adverse selection is an instance of opportunism arising precontractually (chapter 2, section 4, paragraph 4).

try to avoid initiating a large scale dispute over small deviations. In other words, the short term benefits of a major confrontation are too small as compared to the large damage caused. However, such behaviour is expected and taken advantage of[22]. In the long run the accumulated minor damages may or may not exceed the costs of an early termination or renegotiation. A credible termination clause can be an effective deterrent of major defections or of accumulated instances of underperformance, depending on how it is specified. The reciprocity and specific investments often involved in IT provision can deter clients from instituting overly strict termination clauses, thus making space for opportunism. BP managers clearly admitted the opportunistic behaviour of their suppliers, mainly in terms of sticking to the letter of the contract, withholding information and disregarding 'soft' promises (e.g. chapter 4, section 7.1, paragraphs 10, 16, 19 and 21 and section 8, paragraph 9).

Opportunism may also be triggered by contingencies that necessarily accrue due to incomplete contracting. For example, unanticipated changes in demand or user requirements are often treated as work outside the scope of the contract and are charged separately (see chapter 4, section 7, paragraph 9 and section 8, paragraphs 14 onwards). Depending on the nature of this work and on the circumstances, the client may turn such situations to his or her advantage. If such work can be put to tender then the incumbent supplier faces an additional incentive to build good reputation in order to win additional business. The new contract would be the prize for good performance. Often, however, such work is required urgently, the tendering process is costly and time consuming and the incumbent supplier is most likely to enjoy a comparative cost advantage over

---

[22] This is one of the weaknesses of the reputation mechanism as explained in chapter 2, section 7, paragraph 11. Because it is in the interest of both parties to avoid termination, the reputation mechanism induces them to signal to outsiders that the supply relationship is going well, even if there are serious disputes. In other words, either party can behave opportunistically as long as the cost of opportunism is less than the cost of termination. This is more likely, the more specific investment are involved and, hence, the higher the cost of termination.

competing bidders anyway, on account of his position[23]. Such circumstances limit the incentive effect of such a tactic, which would be better suited to relatively large and autonomous projects.

Powerful incentives accrue to the owner of the relevant assets. This is one of the reasons why physical assets are transferred in outsourcing contracts. However, if assets or services are highly idiosyncratic to the users, or if the users are the only residual claimants of operational or investment decisions, then better incentives are created if both asset ownership and responsibility for the relevant decisions is allocated to the users. This argument clarifies the commonly held opinion that 'strategic' services (i.e. services which require significant user-specific knowledge and the gains from which cannot be contractually shared with an external provider) are better kept in-house. In practice, however, the distinction is not always clear. A comparative assessment of the relative importance of each party's decisions, of the skills and knowledge required and of the nature and contractibility of expected gains, is needed in order to allocate ownership and decision making rights[24].

For BP, asset transfers had an instrumental rather than an incentive role (chapter 4, section 6, paragraph 2). Moreover, BP managers preferred to avoid transfers when they involved additional costs either for the physical move or from the additional charges imposed by software vendors. They saw asset transfers as a means of

---

[23] This is the essence of the 'small numbers condition' as defined by Williamson (chapter 2, section 3.2, paragraph 5).

[24] The stylised results of the relevant theories are explained in chapter 2, section 4.2, paragraphs 2-5, where asset ownership is defined as residual risk bearing, and in chapter 2, section 5, paragraphs 4-6, where asset ownership is defined as residual rights of control. In practice, complex structures, including limitations to liability, incentive rewards and penalties and nested or multi-tiered agency relationships, diffuse residual risk bearing. Accordingly, complex provisions are required for the allocation of control rights in order to imbue desirable incentives (see chapter 2, section 5, paragraphs 3 and 10).

increasing the bargaining power of the supplier at contract renewal or in case of termination.

To summarise, this section draws on evidence of supplier opportunism that was observed in the BP case and explores implicit governance mechanisms for countering opportunism. These mechanisms include the informational properties of the broad organisational mode (market, hybrid and hierarchy), reputation, contract termination, competitive bidding and asset ownership. Of these, only information and reputation were found to have significant incentive consequences for BP.

4.2.2. Contractual reward/penalty devices

Formal incentives aim at either encouraging performance (providing positive reward on effort or outcome) or deterring significant deviations from the targets (providing negative reward, i.e. penalty, on effort or outcome). The choice will normally depend on the nature of the service and the way performance measures are defined. For example, network availability can be up to 100%. An encouragement-type incentive would reward the continuous achievement of this target whereas a deterrent incentive would penalise relatively large and/or frequent deviations from the target. The reward and penalty systems implemented in contracts of BP Oil, Air BP and BP Finance are explained in chapter 4, section 7.1, paragraphs 7, 12 and 16 and Figure 4.2.

An encouragement-type incentive is generally preferable when the ideal target is known to be difficult to achieve, when the target itself (e.g. database availability) depends on another contingent factor (e.g. demand for database access) known to be highly variable or in order to foster innovation. Deterrent incentives are implemented to safeguard lower bounds of performance, when a service is critical

to the business and/or when the set target is known to be fairly easily achievable[25]. The assessment of the incentives employed by BP Oil, Air BP and BP Finance in chapter 4, section 7.1, paragraphs 13, 14 and 16 is made on similar terms.

The measure of performance to which an incentive payment is tied, rarely represents the supplier's effort perfectly. There are two sources of imperfection. First, suppliers' performance is not only determined by their own effort but by other exogenous factors as well. Such factors include for example the co-operation of the customers and system breakdowns (variability in the outcome). Second, whether a given outcome (service level or cost) is fully determined by the supplier's effort or not, it may be difficult to measure and verify it accurately (error in measurement).

In such cases, setting up an incentive payment tied to a measure of the achieved result, amounts to shifting an amount of risk to the supplier. Suppliers are willing to undertake risks only if they can spread them over a large number of IT services and/or a large number of customers and if rewards are adequate. Otherwise, it is in the interest of both parties to specify more detailed and accurate measures that eliminate some sources of risk despite the fact that more detailed measures will probably increase the cost of measurement. The incentive system may also account for uncontrollable risk by either allowing deviations from the set target within limits, or by specifying target ranges rather than unitary target values.

In order to reduce the risk that the supplier would have to bear (or alternatively to conserve on the premium that the supplier would charge to bear this risk) the customer may choose to bear some of the risk. This amounts to reducing incentive intensity. In other words, the incentive reward/penalty would be a relatively small

---

[25] In mathematical principal-agent theory, the distinction between rewards and penalties is often irrelevant since, analytically, they are equivalent. In practice, suppliers and clients have to make an informed choice between the two. Some insights are found in the literature on tournaments (chapter 2, section 4.1, paragraph 6).

part of the supplier's total compensation, most of which would be based on cost. Reduced performance may be observed as a result of limited incentive intensity.

If the contingent factors affecting the performance of the supplier have a significant weight in comparison to the supplier's own effort, it may be difficult to design adequate incentives, especially when the customer cannot observe those contingent factors directly[26]. For example, the quality of wide area network services depends on the investments and sophistication of the provider but it also depends on fluctuations in total demand for bandwidth. If unpredictable demand variability becomes a real problem and the provider has no flexibility in bandwidth allocation, then any incentives for speed and quality of data transfers become inoperative as realised performance is effectively left to chance. Fixed price contracts provide excellent incentive for cost reduction but service levels must be stipulated completely and accurately and measured effectively.

In long term relationships, the supplier can be effectively motivated as well as protected from excessive risk if incentives are based on average performance over time, assuming services and SLRs remain unchanged. By allowing a certain number of deviations from the set service level target over time, the customer can forgive the vendor for bad performance in one instance if performance is good in the preceding or succeeding periods. In this way the provider is sufficiently motivated and protected from the effects of adverse contingencies (see chapter 2, section 4.1, paragraph 10). The grief point system examined in chapter 4 (Figure 4.2) operates on a similar basis. As mentioned in section 4.1.3, powerful incentives can also be created by setting a fixed reward (penalty) for performance above (below) a set threshold and reduce the frequency of measurement and/or the amount of information collected. The feasibility of such a measure depends mainly on the

---

[26] In the literature, suppliers (i.e. agents) are usually considered risk averse (chapter 2, section 4.1, paragraph 3). Accordingly, in practice, most outsourcing contracts limit the supplier's liability. The effect of uncertainty (variability in the outcome) and private information (measurement problems) on incentives is examined in chapter 2, section 4.1, paragraphs 4 and 7.

nature of the service and the prevailing conditions of information availability and verifiability.

BP's incentive mechanisms are designed to forgive occasional failures and only penalise consistent under-performance. Moreover, Air BP's contract averages rewards and penalties across services and not just over time.(chapter 4, section 7.1, paragraphs 1 and 7 to 16). However, both allowances create a margin in which the supplier can behave opportunistically without consequences.

Another incentive opportunity arises where two or more services are sequential or somehow interdependent, such as development and maintenance or desktop support and help desks (see chapter 2, section 4.1, paragraph 9). If greater effort in the first service improves the performance (or reduces the costs) of the second, then the contractor can be effectively motivated by undertaking both services. For example, greater effort and performance during the software design stage reduces the cost and improves the outcome of coding. On the contrary, if greater effort in the first service leads to more effort in the second the customer is better off undertaking one of the services himself or assigning it to a different provider. For example, an ambitious proposal during the requirements specification stage will complicate design and development. Alternatively, in anticipation of future work, the provider may specify requirements specific to his capabilities. A system provider performing both tasks is tempted to underperform in the first in order to reduce subsequent costs and complete the project quickly. For the same reason in public procurement, companies carrying out analysis and requirements specification are often excluded from the tender for implementation.

The way service levels are specified has implications for the effectiveness of incentives. Basing service level targets on past performance may give rise to the ratchet effect. In order to avoid this adverse effect, it is recommended that companies commit not to alter their service level targets and incentives, even if they are initially set too low or too high, as long as the IT service provider, the services

and the technology remain the same. Alternatively, the incentive system can be adapted according to past performance without risking the ratchet effect, by simultaneously manipulating one or more of these conditions (see chapter 2, section 4.2, paragraph 6). BP renegotiated and amended some performance metrics during the early part of the contract. At the time of the case study, all metrics were considered stable (chapter 4, section 7.1, paragraph 5).

If a given provider is assigned a number of services, the incentive mechanism should give equal weight to all of them. If some of these services cannot be monitored as effectively as the rest, incentive intensity should be fixed to the lower common denominator permitted by the measurement problems. In other words, if incentives do not place equal weight to all services delivered by the same provider, he will direct all his effort to those which yield greatest returns while maintaining for all the rest the minimum performance permissible by the contract. This has implications for service clustering: services for which incentives are desirable but monitoring is not equally efficacious may have to be clustered separately. Alternatively, additional resources may have to be committed to improving (equalising) the collection of performance information.

The threat of termination can be a low cost means of deterring opportunism (see chapter 2, section 4.1, paragraph 11). It is invoked when significant or sustained defection is observed. It does not require costly third party verification, with the exception of cases that invoke termination on account of perfunctory fulfilment of the supplier's obligations according to the letter of the contract whilst a more positive attitude would be expected. It requires, moreover, governance checks against possible opportunistic use of it by the customer. These checks may entail significant transaction costs, if, for example, the courts or another third party has to ascertain if termination is invoked opportunistically (Muris 1981). BP's contracts make provisions for termination and management is confident that they will be able to manage it without serious damage to the company, if needed (chapter 4, section 7.1, paragraph 2). BPO's grief point system (chapter 4, Figure, 4.2) gives

BPO the right to contemplate termination on grounds of contract breach if SEMA fails consistently below a defined level. However, because of the value of the specific investments committed and the reciprocal dependence that arises, termination is undesirable and, therefore, has diminished incentive effects.

To summarise, this section examines the operation of formal incentive schemes by analysing the incentives implemented in three BP contracts. The distinction between rewarding and penalising incentives, although it is analytically irrelevant, seems to have important implications for practice. The efficacy of incentives is affected by the uncertainty associated with the performance metrics to which the incentive is tied. In this respect, the 'forgiveness' introduced by averaging performance measures for most BP contracts, is noteworthy. Other potential incentive opportunities and pitfalls have been drawn from the literature to complete the discussion in formal incentives.

4.2.3. Other contractual and organisational safeguards

Two more types of incentive instruments are considered here. First, general explicit or implicit policies and guidelines that set the character of the service provision relationship. Second, arrangements that deal with boundary situations and other contingencies. The first kind applies mainly to long term or frequently repeated supply relationships. In the case of internal IT service provision for example, corporate culture and organisational routines may prove more effective performance drivers or deterrents of inappropriate behaviour than formal incentives and penalties. Peer recognition or disapproval are powerful forces guiding the actions of employees. Information sharing is facilitated by formal as well as informal organisational structures and modes of operation.

Framework agreements in outsourcing relationships set, among other things, the organisational and exchange background for service provision to take place. Such agreements serve as the terms of reference for the relationship. Within a

framework agreement, an 'open book' policy and the right to audit serve the need for information sharing (Friedberg and Yarberry 1994), which, in turn, reduces opportunistic potentials (see chapter 2, section 3.4, paragraph 13). Market benchmarking is another means of collecting information and thus introducing an element of competition in both internal organisation and long term contracts. As explained in chapter 2, section 3.1, paragraph 4, such exercises are costly and their informativeness is limited by the extent of service specificity (i.e. the availability of comparable services in the market). For these reasons they are undertaken occasionally and/or when there are other contingent reasons for seeking further information. BP's contracts included all these provisions but they are rarely used (chapter 4, section 7, paragraphs 1 and 5 and section 7.1, paragraph 6).

The cultural match between provider and customer becomes crucial in two main instances. First, many IT services are provided by the contractor's staff within the customer's premises and in direct collaboration with the customer's personnel[27]. Second, various joint decision making structures and processes are in place as explained further on. In such cases the work practices of both parties need to share the same values and objectives. In the absence of complete contracting, corporate culture determines general principles according to which contingencies are dealt with. In long term contracts, compatible corporate cultures between the parties help reduce the costs of bargaining over adaptations during the course of the contract. Culture is, therefore, as important as formal contractual specifications[28].

---

[27] Chapter 2, section 2.1, paragraph 5 stresses the importance of personal (social and psychological) relations in economic exchange. Such relations become even more important when contractor staff work in the client's premises and with the client's staff, although, formally, they report to separate hierarchies.

[28] Culture can be conceived as a set of principles that guide decision making in circumstances that are not sufficiently dealt with by previously agreed rules of conduct. If someone who is aware of a given set of such principles (culture) is able to make informed decisions independently, then culture helps counter contract incompleteness and economises on coordination costs (chapter 2, section 7, mainly paragraph 2).

The prospect of a convivial resolution of future contingencies creates further incentives for the provider to invest on existing and innovative operations.

The relevance and interpretation of corporate culture within BP Oil is evident in at least three instances. First, when criticism was raised on the personal relations of one Partner Resource Manager and the implications this had for service management and user satisfaction (chapter 4, section 7.2, paragraph 4). Second, in the way BP management understands the notion of partnership, not as a loose, open ended, agreement (as it is portrayed in some parts of the literature) but as a closely knit exchange relationship based on sound contractual foundations (chapter 4, section 8, paragraph 5). Finally, the role of culture becomes evident in the dissatisfaction of end-users who lost the personal associations they had developed with the desktop support personnel (chapter 4, section 8, paragraph 9).

In general, the effective incentive system of the governance structure of IT service provision must contain an efficient mix of all the previously described instruments. It may be desirable to introduce a general incentive mechanism that aggregates all service level requirements as well as other (cost related or even qualitative such as user surveys) performance criteria. This is how the reward and penalty mechanisms of BP Oil and Air BP examined in chapter 4, section 7.1 operate. Such a mechanism can serve at least four purposes. First, it integrates local (for individual services) incentives and global (company-wide) coordination. Second, it sets global tolerance limits of bad performance thus controlling risk allocation and facilitating uninterrupted service provision and clearly managed dispute resolution mechanisms. Third, it enables easier top level monitoring and management of the provision of IT, in addition to the detailed measurement of individual service level requirements. Finally, if measurement is not equally efficacious among services, aggregation may lift the problem of non-attention by the provider.

To summarise, this section recognises that formal incentive schemes are limited to the performance metrics to which they are attached. Other aspects of performance

that cannot be measured my thus be ignored. However, a complex outsourcing agreement involves many more expectations that cannot be defined on paper in advance. Formal provisions such as audit and benchmarking rights give the opportunity to BP to restore information parity if the vendor is suspected of misrepresentation. Most of the time, however, these provisions operate as the client's threat position, which deters the supplier from acting opportunistically. Indeed, BP seldom makes use of these rights. Informal mechanisms, such as a climate of trust and a convivial culture between the parties also help to minimise influence costs (haggling) and opportunistic moves when contingencies need to be negotiated. Although these measures, admittedly, do not eliminate opportunism completely, they are important elements of a long term, relational, contract, complementing narrowly targeted performance incentives.

4.3. Organisation and management of service provision

The organisational structures, processes and managerial roles that effect coordination between customer and service provider are the focus of this layer. The managerial roles needed under different organisational structures, the mechanisms by which coordination is effected and the allocation of authority are the three main themes considered.

In the absence of complete contracting, the resolution of contingencies is delegated to hierarchical authority structures. This is the essential feature of relational contracting. With internal organisation, there already exist decision making bodies (individuals or committees) and few additional rules, if any, are needed to specify how various categories of contingencies are to be dealt with. In simple market transactions, the responses to unforeseen contingencies are regulated by the law which specifies rights, responsibilities and appropriate action.

In other contractual long term relations characterised by high specificity, the contract specifies in detail the organisational structures and decision making

processes needed to resolve contingencies. Examples of such mechanisms in outsourcing contracts include the escalation process and the performance review meetings.

Provisions are also made for dealing with boundary conditions such as contract termination. Equivalent arrangements (i.e. firing employees) for the case of internal organisation, are usually left to the discretion of management, of course within the bounds of the law, but the principles governing their operation are similar.

Apart from well-defined contractual clauses, discretionary collaborative behaviours from both the supplier and the client appear to be significant elements of successful procurement relationship and customer satisfaction (Heckman and King, 1994). Such behaviours include various forms of discretionary assistance to the vendor (in the sales process, in the product development process, in the day to day problem solving) which go beyond the letter of the contract[29].

4.3.1. Roles and responsibilities

The degree of managerial involvement in IT service provision varies principally with the type of governance structure. As explained in chapter 2, section 8.1, the more discrete the transaction (i.e. toward the left of the governance continuum of Figure 5.1) the less involved the organisational and managerial arrangements needed. For example, off-the-shelf purchases of hardware and software are handled periodically in a fairly standardised fashion by a small number of people in the customer organisation. Project development contracts require greater managerial attention and involvement to monitor progress and handle changes. Both users and IT staff (e.g. analysts) need to coordinate their efforts with the development team. With fixed term contracts (inside contracting) the contractors are essentially embodied in the organisational hierarchy albeit on different terms as

compared to company employees. Both with internal organisation and with long term contracts, various organisational structures and processes are set up at the interface between customer and provider. The following paragraphs emphasise this case as it is the most complicated.

Typically, at the lower level, the vendor would have some people managing daily service delivery. On the customer side, there would be one or more persons acting as 'user representatives'. These people are assigned to individual services or groups of services and they act as supervisors of everyday operations. In the case of systems development projects, their roles may extend to participation in cooperative project teams.

Above that level, are the people who manage the supply relationship. These managers constitute the main link between users, senior management and the provider. They constitute focal points in the relationship. Their role is to monitor the execution of the contract, and initiate corrective action whenever necessary. Among other responsibilities, this involves managing the incentive mechanisms, collecting and reviewing performance information, following up user complaints and negotiating with the provider(s). They are also the first level at which contingencies and potential disputes are dealt with.

On top of the relationship managers, each side has a managerial hierarchy, its purpose being to integrate the various tasks and activities with the rest of each party's business and deploy the respective authority and bargaining positions when matters cannot be resolved at lower levels. Under normal circumstances these higher level managers intervene less often. Issues referred to those levels include the strategic and long-term evolution of the relationship, contract amendments and renegotiations or extensive contingencies and disputes.

---

[29] See the discussion on relational contracting in chapter 2, section 2.1 and the comments on the role of corporate culture examined in footnotes 27 and 28 above.

The roles outlined above mirror the common denominator of the structures observed at BP, as described in chapter 4, section 7.2. In general terms, these three levels of roles are similar in both internal IT organisation and long term relational contracts. In the former case, the IT supply contract is implicit and is almost continuously adapted and renegotiated between user and IT management.

Another parameter affecting the degree of managerial involvement is the uncertainty associated with a given service or technology. A new, immature technology which is relied upon for significant business benefits, naturally attracts substantial managerial attention until it stabilises and can be reliably delegated to an internal or external support team. Videoconferencing was such a case for BP Oil for a period of time when wide area networks were used for batch data transfers. Internet and related technologies may also attract similar managerial attention in those organisations that are trying to integrate them to the mainstream of their activity.

To summarise, this section reiterates the managerial roles observed in BP. An attempt has been made to place these observations within a broader theoretical context, covering a wider range on the governance continuum.

4.3.2. Co-ordination and communication mechanisms

Communication and coordination requirements of governance structures range from trivial on the left of the governance continuum (Figure 5.1) to very elaborate towards the right hand side. Within long term relational contracts and internal organisation, communication and co-ordination occurs both horizontally across the interface and vertically within the organisational hierarchy of each party. We can usefully conceptualise three broad classes of information exchanges which also correspond to different hierarchical levels. First, daily operational information about service level achievement and respective problems, second, aggregate

information on performance, the operation of the incentive mechanisms and persistent problems and third, long term and strategic considerations.

Steering committees in internal IT organisations and joint boards in outsourcing contracts are equivalent structures providing top level coordination. Such bodies meet at regular intervals to provide top level direction and make corrective adjustments to the ongoing relationship.

It is at these joint bodies and at other interaction points that the managerial style and organisational culture play a crucial role in facilitating the co-operation and resolution of decisions for which it is impossible to make contractual provisions in advance. These are decisions related to adaptations to changing circumstances and occurring contingencies that are unpredictable at the time the contract is written.

The size and complexity of the interface between client and supplier in terms of number of people and contact (interaction) points affects both the organisational costs as well as the effectiveness of IT provision. Whilst necessary structures for the collection and dissemination of information and decision making, they are also sources of bargaining and influence costs[30]. The presence of high asset specificity and/or uncertainty generates a considerable co-ordination burden that requires complex structures and processes. It is expected that as a given governance structure evolves and matures, organisational learning, operational knowledge and mutual trust develop and should enable the simplification of the interface.

---

[30] When there are open channels of communication between decision makers and other parties, and the latter have private information and/or divergent interests, they will dissipate resources trying to influence the decision outcomes to their own private benefit. These are influence costs (chapter 2, section 4.2, paragraph 8). While open communication channels are necessary for informed decision making, appropriate governance mechanisms (including, for example, decentralisation and corporate culture) are needed to mitigate influence costs.

BP would have preferred to contract with one entity, if the European Court allowed it (chapter 4, section 8, paragraph 6). Instead, BP's contracts make one of the suppliers responsible for all problems and failures at any given site, even if they are due to a third party's omissions or actions. This provision is in place so that influence activities and coordination problems might be minimised (chapter 4, section 7.1, paragraph 12 and section 8, paragraph 10). Even so, because of the multitude of different contracts service various business units in one site, the communication and coordination arrangements were found to be excessively complicated. BP and SEMA were looking for ways to simplify the interface (chapter 4, section 7.2, last paragraph).

To summarise, this section addresses the interaction taking place between client and supplier in the course of service provision. Various review meetings and coordination procedures were observed, including some reflections on their effectiveness. An attempt is made to reconsider these issues from a more general perspective. Clearly, the observations made at BP and the present discussion are largely limited to long term service contracts, as opposed to other forms of governance such as requirements contracting. Having said that, long term service contracts provide a good demonstration of their relationality and their similarity to internal organisation.

4.3.3. Locus and scope of decision making

Decision authority in discreet transactions is well demarcated and the courts typically handle any exceptions. In intermediate and relational governance structures the allocation of decision rights to the various managers and co-ordination bodies will have to depend on the nature of the services in terms of asset specificity[31], on the importance of each party's decisions and investments for

---

[31] Transaction cost economics relates asset specificity, as a dimension of transactions, with adaptive capacity, as a dimension of governance structures. Greater asset specificity requires, among others, governance structures with the capacity to adapt cooperatively (as opposed to autonomously).

service provision[32] and on any benefits of specialisation in knowledge and information[33].

Four authority roles are identified by Galbraith (1977), in decreasing order of authority power, namely, responsibility, approval, consultation and informing. The allocation of each of these types of authority may have to be specified for every task or decision, depending on the complexity of the service, of the governance and of the supplier-customer interface. They may be allocated to the same or different persons or organisational bodies (e.g. committees). Purely technical decisions would be the responsibility of the provider. Major technological or organisational decisions may require the approval of the client organisation. The provider would also have an informational consulting role on user-specific issues.

The juxtaposition of the four stages of decision making (Jensen and Meckling 1983, see also chapter 2, section 4.2), namely initiation, ratification, implementation and monitoring is useful in putting matters into perspective. With both internal and external service provision, all four stages are, in general, shared by users and the provider. Either users or the provider may initiate a change or a new development and both parties have to approve the business and technical implications respectively of the decision before it is implemented. Only certain supply-side decisions are decided upon by the provider alone. The implementation of most decisions has both technical and organisational aspects. Progress and performance

---

Cooperative adaptation mechanisms include joint decision making structures and processes for private dispute resolution. See chapter 2, section 3.3, mainly paragraph 4 and section 3.4, paragraphs 8-9.

[32] Reference is made here to positive agency theory and the theory of incomplete contracts, which relate asset ownership with decision rights for incentive purposes. See chapter 2, section 4.2, paragraph 4 and section 5, paragraph 4.

[33] As explained in chapter 2, section 6, mainly paragraphs 5-6, the collection and processing of information and knowledge is a costly activity. Appropriate distribution of monitoring and decision making authority (e.g. central or distributed control) can improve the efficiency of the process.

monitoring is also a joint responsibility, depending on the respective costs. In general though, as noted above, more technically oriented decisions are in the area of responsibility of the provider while business specific decisions are handled by users.

However, such distinctions are not always clear. Not only is the nature of decisions inherently mixed but when the IT department is a profit centre (or when there is an external supplier), the two parties also share residual claims (the net benefits of the implemented decisions). Where cost effective measures of the interface between the business and the technical side are available, decision authority and responsibility can be clearly allocated by a nearly discrete contract. The contract will stipulate deliverables and performance. Where such separation is not possible, the contract (if there is one) would have to make allowances for joint decision making. BP's contracts (the framework agreement) make explicit provisions for renegotiation and adaptation of services and service level requirements and for the resolution of disputes (chapter 4, section 7, paragraphs 3 and 11). These provisions are essentially decision making processes, involving staff from both parties.

The decision making structures instituted as part of the governance must have clear objectives and authorities. The problems that BP faced with the operation of the Outsourcing Review Board (chapter 4, section 7.2, paragraph 6) show that all the elements of governance, including decision making bodies, must be justified on grounds of effectiveness and/or efficiency. In other words, such structures must either add value (e.g. by generating improvements in services and their management) or mitigate transaction costs (e.g. alleviate information asymmetry through monitoring and incentives).

The system of service level measurement described in previous sections generates a large amount of information which can potentially become valuable in improving both the organisation and management of service provision as well as the services themselves. Appropriate communication channels are needed to collect and feed all

this information to the decision making bodies of both parties for continuous process and service improvement. Dedicated reviews may also be instituted. This is how service level requirements and performance monitoring can create value as discussed in section 4.1.2. BP explicitly recognised the need to improve the usefulness of the performance reports, beyond their instrumental role in the reward and penalty (grief point) mechanisms (chapter 4, section 7.1, paragraph 6 and section 8, paragraph 12).

To summarise, this section attempts to combine theoretical insights with case study evidence on the allocation of decision making authority for contract management. It is interesting to note that the stylised conditions studied in formal economic models do not apply directly in practice, where a much higher degree of complexity is observed. In particular, we find that in practice, residual claims do not rest solely with one or the other party. Inasmuch as compensation is fixed and supplemented with formal rewards and penalties, the supplier is the residual claimant. Simultaneously though, there are price adjustment mechanisms addressing possible changes, and there are limited liability clauses protecting the supplier. These provisions render the client residual claimant. As a result, the allocation of decision management and decision control is not straightforward, as suggested by the positive theory of agency. The actual allocation of decision rights depends on the substantive nature of the decisions themselves and on additional factors such as information costs, influence costs and the potential to create value-added from these decisions. BP shared decision management and decision control tasks with the suppliers in different ways for different tasks.

5. Conclusions

This chapter went into considerable detail to elaborate a compact idea: the governance of IT provision should be established by considering a number of contractual, organisational and other elements in conjunction with the characteristics of the services. Because of the variety and inter-relationships among

technologies, systems and services, they should be clustered in such a way that the different governance structures assigned to each cluster strike a balance between specialising the governance to the characteristics of services and taking advantage of synergies (minimising coordination cost). These recommendations are arrived at by assessing the relevance and applicability of the theories reviewed in chapter 2 to the contracts observed during the case study at BP (chapter 4).

As requirements, technologies, perceptions and market conditions change over time, companies strive to maintain lowest costs and highest quality of services. This requires a degree of flexibility in governance. Flexibility in the organisation of IT provision can be considered at two levels. First, internal flexibility lies within a given governance structure. When a change or contingency occurs, the existing arrangements may be able to accommodate it. With a commodity product which has few learning or lock-in effects, the market will adjust prices and product specifications. Such an example might be the supply of PCs. Upgrades and price reductions in the market are readily transferred to the client through the same supply contract with limited or no negotiations (autonomous adaptation, chapter 2, section 3.3, paragraph 4). The same simple type of supply arrangement cannot provide similar flexibility when complex services are involved. Bespoke software development cannot be provided through a similar contract as PC supply. Changes in requirement or technical obstacles cannot be dealt with because there is no market price neither for the software itself nor for changes and obstacles. More involved structures are needed wherein management discretion and intervention are employed to reorganise and reschedule the project (cooperative adaptation, chapter 2, section 3.3, paragraph 4).

Second, external flexibility lies in the ability to switch governance structure or supplier altogether. When internal flexibility is insufficient (i.e. when the self-enforcing range of the contract is exhausted, chapter 2, section 1, paragraph 13), companies must be able to switch to another form of governance and/or to another supplier. External flexibility is an inverse measure of the cost of switching.

Changing PC supplier is relatively easy because contracts are straightforward and the product is standard. Changing software developer amidst development is next to impossible because the product and to a lesser extent the technical knowledge are specific to the customer and the supplier respectively. Vertical or lateral integration can be a result of the adaptive failure of such a contract.

The more long-term, complex and specific the governance of IT provision, the less its external flexibility. In such cases though there is usually more scope for internal flexibility. The two are, up to a certain extent, interchangeable. For example, it is very costly to terminate and replace a long term total outsourcing agreement because the two parties are closely tied to each other: the supplier has made specific investments for the particular contract and the customer essentially relies its business on the particular skills the incumbent supplier has developed. External flexibility is therefore limited. Such agreements however include (or should include) complex provisions for dealing with uncertainty via managerial intervention or third-party arbitration without reaching termination. Such mechanisms ensure internal flexibility.

The application of S-CAGE (or more generally the management of the governance portfolio) should be seen as an ongoing process rather than a one-off exercise. The framework provides the means both for making continuous assessment as well as for achieving the flexibility required to implement adjustments.

The S-CAGE framework provides a basis for organising the relevant economic literature. Table 5.1 summarises the correspondence of the theories of chapter 2 with the three layers of governance in S-CAGE (Figure 5.4). Based on the BP experience, the framework also proposes a strategy for tackling the problem of IT governance in practice and highlights a number of particular issues that need to be taken into account. Further evidence of the usefulness of S-CAGE in analysing and understanding practical issues is presented in the following chapter.

CHAPTER SIX

Using S-CAGE to Understand IT Outsourcing Practices

1. Introduction

The S-CAGE framework was introduced in the previous chapter as a normative framework for setting up the governance of IT service provision. As such, it can also be used as a means of analysing, understanding and communicating the governance of IT provision in a given setting with the aim of assessing and, possibly, adjusting some of its elements. This chapter presents a description and assessment of IT outsourcing practices in two companies, namely, ICI and Anglian Water, in order to demonstrate and evaluate the use of the framework as a descriptive and analytical instrument. In each company the relevant information has been collected by interviewing a key individual and reviewing relevant documents. An open-ended interview guide was prepared for use in both cases (see appendix). The interview guide follows the structure of the three layers and sub-layers of governance presented in chapter 5, section 4. The issues raised and the questions asked reflect the arguments that are elaborated in chapter 5, sections 4.1 to 4.3.

Unless it is employed in a new start-up company where IT is introduced anew, any attempt to use S-CAGE would have to start from an assessment of the arrangements already in place. This is how the two cases that follow have been approached. As a first step, the mapping of the existing governance portfolio reveals the main policy followed and the implicit clustering of services that has been established. Each cluster (contract) can then be subject to analysis of its governance details, which, in turn, can subsequently provide a basis for changes.

The methodological basis for these case studies has been discussed in detail in chapter 3. Briefly, their purpose is to assess the theoretical power of the S-CAGE framework in alternative contexts, other than that in BP which informed its formation. This chapter does not constitute a formal test of the proposed framework in the mainstream positivist sense, for two main reasons. First, although the framework has been presented in a normative way, it is not tested in a complete reorganisation of IT provision. This would entail serious practical difficulties. Second, it does not cover a broad sample of organisations and contracts. Both contracts examined are for data centre services. Therefore, there is ample scope for trying out the framework in many different kinds of contracts along the governance continuum (including internal organisation), for different kinds of services.

What these case studies achieve, is to demonstrate the S-CAGE framework in action, in organising, interpreting and assessing the multitude of elements comprising the governance of outsourcing arrangements. Outsourcing contracts are lengthy and complex documents where the management implications of their various clauses are confounded by the enormous detail and esoteric jargon of their style. At the very least, S-CAGE is a framework for organising and understanding all the elements of governance in a relevant manner, from a management perspective. However, the value added is in interpreting and assessing the elements of governance (contractual provisions and management processes) and in highlighting their implications (and possible alternative configurations), on the basis of pertinent economic theories.

At the outset, the two companies examined here have adopted outsourcing for the same purpose, namely to facilitate the transition from a mainframe to a client-server computing environment. The case evidence corroborates several similarities expected in the structure of contracts. For example, the structure of compensation and the formal incentive systems are very similar. However, there are differences that merit closer examination. The most important difference is in the service clustering strategy adopted by each company.

In the following sections, the cases of ICI and Anglian Water are considered in turn. The presentation of each case follows the structure of the S-CAGE framework. Some general background information on the company and their IT outsourcing practices is followed by a description of the respective governance portfolio. The discussion of the governance portfolio reveals the implicit clustering of services and the reasons/causes of its pattern. One outsourcing contract is subsequently selected for an analysis of its governance details. The governance details are categorised following the three layers of governance (Figure 5.4). A critical commentary on the findings completes each case. Finally, the concluding section outlines the lessons learned and draws some implications for the usefulness of the framework.

2. Corporate IT Outsourcing Practices At ICI

Imperial Chemical Industries Plc (ICI) is one of the largest companies researching, manufacturing and selling paints, materials, explosives and industrial chemicals. It has operations in over 15 countries world-wide with a total group turnover of £10.5 billion in 1996. Over the past two years, the Group has been implementing major restructuring projects involving divestiture of certain businesses and acquisition of others. Broadly, their aim has been to shift the company's portfolio away from cyclical industrial chemicals and towards market segments with high growth potential, namely specialty chemicals addressing specific customer needs.

This intense restructuring activity, which made regular headlines in the business press, has limited the access to IT management for the purposes of this research. The main contact has been made with the Group IT Manager. Consequently, the emphasis in this discussion is from the corporate point of view. A contract document for data centre services was also made available.

2.1. Strategic direction

The strategic direction for IT in ICI stems from the business strategy, which defines key business processes and a set of strategic capabilities for each process. IT is considered almost a commodity and is expected to support and facilitate these capabilities. An essential skill for ICI, developed and maintained in-house, is Information Management. With this term ICI is referring to a small cadre of approximately 100 people across the Group with both technical and business expertise who can blend business requirements and technical opportunities. Their skills are critical in, among other processes, the context of writing and managing IT procurement contracts and particularly service level agreements.

For different ICI businesses, there have been various imperatives for outsourcing. Overall, none of ICI systems or services are perceived as critical or core to the organisation. As a result, the directive at a strategic level is that ICI does not want to maintain a capability for delivering technical services and therefore all IT systems and services are candidates for outsourcing. It then becomes a question of what can the market provide and under what conditions. Indeed, for the Group IT Manager the primary driver for outsourcing is the marketplace, the capabilities offered by vendors and the conditions of competition. ICI's present IT governance portfolio appears to be fairly fragmented and ICI faces the burden of integrating and coordinating suppliers. ICI envisages the possibility of being able to procure whole IT 'processes' from a competitive market in the future, although it is recognised that the idiosyncratic nature of systems and services, as well as the structure of the market, impedes such approaches at present.

Outsourcing exercises at the corporate level are led by a steering group, headed by the Group IT Manager. The steering group directs a number of professionals who look after different aspects of the process. Four categories of management skills are important for ICI in the context of outsourcing. First, commercial purchasing and, second, negotiation and legal skills are typically found to be serving various business activities within the organisation. Third, information management skills

and, fourth, contract management skills are being developed among IT managers within the company. It is clear to ICI that the management effort involved in setting up and managing service contracts varies from one service to another. For example, wide area networks are ranked first, data centres second and new system developments third, in increasing order of complexity and management effort. This ranking can be interpreted to correspond to an increasing order of asset specificity involved in each case. Although this terminology is not being used at ICI and despite the perception of IT as commodity, the pragmatic implications of specificity are clearly understood and acted upon.

Examples of other approaches to outsourcing within the ICI group include ICI Paints in Europe. They have contracted with CSC for consultancy, implementation and management services for the transition from legacy systems to a distributed architecture. The main focus has been on the transition process. For the purposes of this contract, CSC undertook the operation of existing mid-range systems and parts of the management of existing contracts with Origin (for access to the Data Centres) and other suppliers. CSC also takes full responsibility for the new systems and applications.

ICI Films in North America contracted with EDS for a similar set of services (running existing mid-range systems, consulting and implementation of new developments). However, the driver in this case appears to have been a dissatisfaction with existing operations and the desire for fresh thinking, new skills and new value adding opportunities from IT. For this purpose a member of the EDS team sits in on the IS strategy meetings of the business. This practice comes in sharp contrast with the opinion held at the corporate centre in London, that suppliers of systems and services should not be involved in decision making. However, by forging such a collaboration, EDS is more likely to deliver business-specific, value-adding innovations from which suppliers are deterred under the 'arms-length' relationships in more usual outsourcing contracts (see relevant discussion in chapter 5, section 4.2 and chapter 7, section 5). Further details on

this contract were not available to make possible a better assessment of the significance of EDS's participation in ICI North America's IT strategy meetings.

2.2. Governance portfolio

At the corporate level, ICI is managing several IT contracts. Their overall governance portfolio across the organisation appears to be fragmented, as they procure data centre services, wide area networks, help desks, desktops, servers, software and individual projects separately. One reason for not having everything under a single large contract is that in this way they try to maintain competitive procurement deals. Another stated reason is that the capabilities of suppliers to provide integrated services are not considered to be adequate at present. It also appears to be the case that as ICI is a diversified group, individual businesses have a fairly high degree of autonomy in selecting their own technical standards and procurement policies. This fragmented structure has, admittedly, been a source of concern. ICI is responsible for ensuring service integration and supplier coordination. In most cases, one supplier is 'given the lead' to handle integration and coordination. A similar practice was also observed at BP (chapter 4) where one of the suppliers was being held responsible for all operations at a given site, although some services were provided by others. Alternatively, a third party may be hired with the mandate to manage the interface between ICI and the existing IT supply chain. Such practices are employed to economise on coordination costs, an issue discussed in the context of clustering in chapter 5. The main outsourcing contracts under the control of the corporate centre are described below.

Wide area networks (WAN) are provided under three contracts, covering Europe, US and Asia/Australia respectively. Approximately 20 people world-wide are responsible for managing both the commercial aspects of the contracts as well as the integration of technical platforms and suppliers. Of particular concern to them is the provision of Lotus Notes operability across their global networks. WAN contracts have 2-3 years duration and ICI are constantly seeking competitive prices in the market. Appropriate pricing clauses have been included in the respective

contracts, which, in general, specify the right of ICI to market-test the prices offered by incumbent suppliers and seek adjustments.

Desktop support and help desks are organised in a variety of ways across the group. Local area networks and help desks in Europe, one level below the corporate centre, are organised under a single contract. The next outsourcing project being planned at the time of the interviews would include high level servers, comprising software and equipment lying between WAN, data centres and the desktop.

Data centre operations shared by a number of ICI businesses in Europe have been contracted out to the Dutch company Origin. Origin has been given the option to bid competitively for new developments on distributed systems and has been successful in winning some of them. By the time the contract expires in 2001, the existing mainframe operations are expected to have been phased out. This transition is implemented in a decentralised way, with each ICI business unit making its own decisions.

The phasing out of mainframes was the main driver for outsourcing in this instance. Among other suppliers, Origin was interested in undertaking this contract profitably as they were in the process of expanding their presence in the UK and the large data centre operations of ICI (including staff) offered a suitable opportunity. In contrast, other suppliers were found not to be interested in increasing their stock of data centre capacity and would have absorbed the costs of closure (redundancies etc.) for a higher premium. ICI perceive their general policy to outsource in this case, as the ability to grasp a market opportunity. Having said that, ICI would have anyway outsourced this service on the basis of their strategic decision to purchase, rather than maintain internally, technical capabilities. The governance elements of this contract are discussed in the following section.

2.3. Origin contract: Governance structure

The following analysis is based on the contract document and on related comments from the Group IT manager. The structure of the sections below mirrors the three layers of the S-CAGE framework (chapter 5, Figure 5.4) and the discussion follows part B of the interview guide in the appendix. The interview guide is a device for seeking information and subjecting it to the inspection and assessment of S-CAGE. As such, the interview guide was used both for conducting the interviews as well as for extracting relevant information out of the contract, which, like most contracts, is a complex and esoteric document. Following the interview guide, relevant information from the contract and the interviews is presented and commented upon, based on the respective arguments in chapter 5.

2.3.1. Assurance of Service Delivery

2.3.1.1. Service specification

The contract was entered into in December 1995 for a 5-year period. This contract defines a number of services relating to the operation and maintenance of ICI's mainframes and the software running on them. It gives Origin exclusive right to provide these services. Although ICI asks Origin to provide a number of related and additional services, it reserves the right to invite tenders for such services from other suppliers and include Origin in the process.

Basic services are distinguished from additional services. Basic services include applications maintenance and support, application enhancement, computing platform services (hardware, systems software etc.) and other miscellaneous services. A series of detailed activities are defined under each category of basic services. Most of these services and activities relate to a number of software products, each of which supports one or more applications. Additional services include various development, enhancement or other change projects that were under way at the time of contracting, as well as similar services that ICI requests from time to time on an ad-hoc basis. The contract specifies that Origin should

have available resources to support services such as systems operations (mainly including applications support, and facilities management), development and installation of new applications and IT consulting (on IT strategy, BPR etc.), upon ICI's request (and according to defined contract and service change management procedures).

The specificity of these services is directly manifest in the contract in two main ways. First, the different software owned by ICI at the time of contracting is classified as exclusive, restricted and unrestricted, according to the permitted usage, whether by licence (where the software is licensed) or otherwise. Exclusive software is for the exclusive use of ICI and its subsidiaries, restricted software can also serve other predefined companies and unrestricted software can be used to serve the needs of any Origin customer. 64% of the software packages listed are for exclusive use and 30% are unrestricted. Although the mix of human and technical resources required for the exclusive software may not be unique in its entirety, the majority percentage of such software indicates a significant degree of specificity.

Second, the contract explicitly recognises the importance of certain key personnel who are essential for the proper provision of the services. Origin is asked to retain that staff for a period of six months into the contract to support the services. In particular, key personnel include 19 staff, mainly functional analysts and team leaders/managers.

In addition to the above two explicit manifestations of service specificity, throughout the contract, four further elements implying the existence of asset specificity are identified among these provisions. First, controlled access to premises (site specificity), second, minimum threshold of payments (a form of 'hostage'[1]), third, intangible assets and know-how (human asset specificity) and

---

[1] See chapter 2, section 3.4, paragraphs 11-12 and chapter 5, section 4.2, footnote 19 and accompanying text.

finally, segregation and disentangling of networks in cases of sale of business units (physical asset specificity).

A final indication of asset specificity is a provision that prohibits either party to employ any of the other party's staff. Anecdotal evidence suggests that as certain client or vendor employees accumulate knowledge of the other party's business during the course of service provision, they become indispensable to that other party (human asset specificity), which often attempts to employ them. Whereas the accumulation of such knowledge is desirable and important for purposes of effectiveness and efficiency, the prospect of such shift of employment induces the original employer to discourage knowledge accumulation by means of job rotation and other tactics. Appropriate contractual clauses such as the one reviewed here restore incentives and mitigate opportunism.

In terms of uncertainty, data centre operations are generally considered fairly stable services. Computer processing volumes (CPU times, disk storage etc.) for basic services and man-days for additional services were forecasted at the time of contracting for each of the five years of the duration of the contract. On the basis of these forecasts, ICI was bound to pay for a minimum threshold of consumption in each year. This demonstrates low perceived uncertainty for these services.

Having said that, the major source of uncertainty for IT services at ICI during the last few years has been the intense restructuring activity at corporate level. As mentioned in section 2.1 above, ICI has been divesting business units and acquiring other businesses in an effort to reorient its portfolio. Since data centre operations serve many business units within the Group, such restructuring has had a profound effect on the nature and quantity of demand for IT services. The contract makes a series of provisions for managing and adjusting to such changes and a recent amendment to the contract (in February 1998) revised, among others, the forecasted quantities for each of the remaining years as a result of corporate restructuring.

A lengthy set of provisions specifies how changes of any kind (in services or in the terms of the contract) that are not explicitly considered in the agreement, are to be initiated, negotiated, agreed, documented and implemented. Of course, trivial changes that are part of the process of delivering defined services (such as variations in computer processing capacity volumes) are excluded from the specified procedure and the contract goes to significant lengths in defining what constitutes a (non-trivial) change. Changes to services are classified in three main categories, in increasing order of the extent of change. A fourth category concerns amendments to any of the other three, once the latter are under way. For each category, a process for authorising, managing and documenting the change is specified. Furthermore, a list of special cases as well as pricing guidelines are also supplied. These procedures are under the control of the management teams of both parties (see section 2.3.3) and have important incentive implications (see section 2.3.2).

In terms of transaction cost economics, the change management procedures serve as an adaptive mechanism (see chapter 2, section 3.3, paragraph 4). Other adaptive mechanisms in ICI's contract with Origin include provisions for benchmarking and indexing service charges, the dispute resolution procedure and numerous minor clauses which defer decisions to a later point in time, during contract execution. The management structure established in the contract (see section 2.3.3), is another important mechanism for adaptation to circumstances that arise during contract execution and that cannot be anticipated at the time of contracting.

As discussed in section 2.1 above, an important rationale of the decision to outsource data centres was to take advantage of the scale economies achievable by IT suppliers, particularly in view of the fact that ICI had made the decision to phase out the mainframes gradually over the five-year period of the contract. Clearly, it would not be cost-effective for ICI to maintain a data centre operation for constantly declining volumes and then having to bear the costs of closure as well. Origin needed additional mainframe capacity to establish and expand its presence

in the UK. Scale economies and efficient resource utilisation, therefore, were important factors in this outsourcing decision.

Resources shared among the services covered by this contract include the common hardware platforms, the software suites and the specific sites where the equipment is installed. Although not explicitly stated, the skills associated with the services are to a significant extent the same and generic (e.g. computer-platform operations and software enhancements) and are only specialised where particular hardware or software products are concerned and where ICI-specific knowledge of the business is required. Whereas the common generic skills mentioned above are, in principle, readily available in the market and divisible (e.g. a company can readily contract for a set amount of programming man-hours), it is the possibility of scope economies in terms of specialised skills (particularly ICI-specific) that matters most (see also chapter 2, section 3.4, paragraph 5).

The interfaces among different software or systems constitute a significant part of each product's description. Apart from such descriptions the contract does not explicitly acknowledge any performance implications of inter-dependencies between systems and services. Having said that, express provisions are made with respect to the obligations of Origin to cooperate with any other suppliers who may provide IT services to ICI and where these services interface with Origin-provided services. In particular, Origin's cooperation is sought in three ways. First, to provide access to information, premises and networks as needed. Second, to ensure that systems and services are well integrated. Third, to secure the resolution and prevention of problems in service quality that may be attributed to the other supplier's operations. Although these obligations are restricted to Origin's side of the interface with any other service supplier, the contract seeks the positive collaboration of Origin in matters relating to interdependencies with services outside Origin's control.

For most part, the specification of systems and services refers to the material conditions of *how* to deliver the services, as opposed to *what* must be delivered (in

terms of business processes). To begin with, services are charged on the basis of resource usage (in terms of CPU time, data storage volume, and man-days). The hardware and software to be used is also specified in detail. The measures employed for purposes of calculating performance incentives (penalties) are availability, reliability, response and serviceability. In comparative terms, these are the most user-oriented (non-technical) requirements set in the contract. Services are defined in terms of a series of activities, which, in turn, describe in fair detail how things should be carried out. For example, application maintenance and support services comprise eight activities (e.g. *release management*) and numerous sub-activities (e.g. *handover acceptance* is defined as "*formal acceptance of maintenance & support responsibility concerning error modifications and/or enhancements by the basic Applications Maintenance and Support Services team*").

In addition, the contract makes fairly extensive requirements for disaster recovery, implementation of the changeover, internal audit, project management and quality assurance. In all these respects, the contract is not limited to outcome requirements, but involves ICI extensively in the decisions as to how the respective processes are to be carried out. Although the respective provisions are typically extensive, in most cases the contract does not address the matter in its entirety but defers the decisions on particular plans of action to a later date. Origin is asked to draft the plans, subject to ICI's approval. These, among other, mechanisms clearly reflect the relationality of the contract (see chapter 2, section 2.1). Moreover, they reflect the measurement problems associated with defining requirements such as disaster recovery and quality assurance only in terms of outcomes and without reference to process. The contract repeatedly makes reference to best industry standards (which may be well defined) and draws on the judgement of third party experts in some cases of disagreement. These and other similar provisions, which make reference to results rather than process, do not appear to be adequate to cover all the requirements of ICI. Reference to process (*how* things are done) seems to be necessary.

Throughout the contract ICI maintains effective discretion and control on what Origin should do and how. This discretion is primarily manifested in terms of approval of proposals, plans and reviews submitted by Origin and also in terms of initiating orders for additional or modified work. More explicitly, the contract specifies *"[...] that ICI may from time to time issue directions [...] ICI shall be entitled to require compliance with the directions as soon as reasonably practical"* even if the relevant documents (orders) have not been formally agreed upon. This clause appears in the context of Origin's responsibilities for interfaces with other suppliers or with activities retained by ICI, but it is representative of the discretion retained by ICI.

The performance incentive (penalty) scheme gives an indication of the criticality of services. For the purposes of performance measurement, each service level requirement for mainframe operations is assigned a weight. In aggregate, the response, availability, reliability and serviceability of on-line processing, batch processing and help desk support services are allocated the vast majority of weights (nearly 90%). For comparison, the rest of the weights are allocated to other related services including output handling, backups, disaster recovery, access control and contract management. In other words, the availability and response of mainframe applications is what matters most for ICI.

2.3.1.2. Service level requirements specification

As far as service level requirements for inputs are concerned, the contract does not specify the amount of human resources required for any service, with the exception of key personnel (see above), man-days or man-years for additional projects and implicitly by defining responsibilities for contract management. However, for purposes of pricing services, detailed marginal rates are defined for all resources, including human resources (where rates vary with rank) and computer processing resources (including CPU time, data storage volume etc.). Detailed provisions are made for costing and payment calculations and procedures.

Agreed rates can, over time, be adjusted according to market indices and benchmarking.

ICI reserves the right to ratify all costs that are not anticipated by the agreement, before they are committed. Such costs include those arising from additional services and changes to services (covered by the change management procedures), as well as from items that may not be included in the agreed figures.

As far as output measures and targets are concerned, the contract makes detailed specifications. For each computer platform, a series of service activities are defined and for each activity, one or more measures are to be applied to each software product. Where relevant, measures may include response, reliability, availability, serviceability and service hours. For each measure in each activity the following details are supplied: definition, required level (target), commitment period, measurement responsibility, reporting frequency, point of delivery, measurement method and other comments. In addition, user satisfaction surveys are carried out every six months.

The service level measures that were in place while the data centre was operated in house and the average performance levels achieved by ICI have been incorporated into the contract. ICI and Origin agreed on a 'service assessment period' of at least three months during which to evaluate and, if necessary, revise performance measures, targets and reporting methods. The contract directs the parties to establish revised service levels which bear the same relation to the service levels specified in the agreement "*as the actual performance of that service over the service assessment period (as agreed between the parties) bears to the performance which is recorded in the reports produced [...] in respect of that service during the service assessment period*". In other words, because service levels in the original contract were set on the basis of reported performance, if the extant reports under-represented actual performance then service levels should be revised upwards and *vice versa*. This paragraph attempts to enable a fair and efficient revision of

performance targets without triggering adverse incentive consequences[2]. It is not clear how effective is this clause in doing so, but no significant problems have been reported in this respect.

In another paragraph, the parties are directed to *"seek to agree formats and measurement criteria for reports which clearly and objectively demonstrate for each of the services to which such reports relate, whether the required service levels have been attained in respect of such service"*. This paragraph attempts to ensure the completeness and precision of service level requirements (see chapter 5, section 4.1.2). In other words, the contract emphasises the need to devise measures of performance that are representative of actual performance and unambiguously measured so that any possibility for opportunistic misrepresentation of actual performance may be avoided. In the same clause, furthermore, the contract urges the parties to avoid opportunistic manipulation of actual or reported performance during the service assessment period and refers any disagreements to the dispute resolution procedure.

The specifications of service level requirements include further evidence of completeness and precision. In order to cover different facets of a service, more than one metrics are introduced (e.g. response, availability and reliability for on-line processing). In order to capture all aspects of performance for any one metric, more than one targets are set. For example, the response metric for on-line processing is associated with three interval targets (percentages of transactions to be completed within set time periods) and one overall average response time.

The Group IT Manager is envisaging the possibility of linking the assessment of IT suppliers to overall business performance indicators but this is not considered attainable at present. The system of service level measurements is considered to be satisfactory and fairly comprehensive.

---

[2] See chapter 5, section 4.2.2, paragraph 10.

2.3.1.3. Service level measurement and reporting

Service level requirements are measured either manually or with the use of measurement equipment and/or software. Service level targets apply to monthly, quarterly or annual periods. Most measures are reported monthly with few exceptions. Origin has responsibility for measuring and reporting the majority of service level requirements. Requirements pertaining to contract management (e.g. attendance to meetings, delivery of reports and invoices etc.) are the exception and are monitored by ICI.

Some measures are not constantly being monitored/measured but either failures are flagged by customer complaints or a customer complaint initiates a measurement and review period. Those service level requirements which are not regularly measured include either trivial requirements that are easily met or cases where failures are expected to be rare (e.g. security breaches). Either way, this practice serves to minimise the administrative burden (cost) of measuring and reporting. None of the metrics is measured at random.

Some service level requirements are included in the calculations of performance incentives (penalties) but not all. Eleven out of 29 metrics are excluded from the performance index. These requirements constitute contractual obligations for Origin but most of them do not have any financial incentive implications. Metrics representing outages (service stoppages) are linked to financial penalties, which are, however, calculated in a different way (see section 2.3.2 below). Some metrics without financial implications are trivial (e.g. *"business communication directory changes will be completed within two working days"*). Others may not be trivial but the achievement of the target may be hampered by significant uncertainty (e.g. all metrics related to the response and accuracy of estimates and quotations for enhancements and additional projects). Such service activities and measures could be candidates for encouragement-type incentives (rewards for high performance rather than penalties for low performance) but such incentives are totally absent from this contract. The extent of financial implications of performance shortfalls

with respect to any metric is also an indication of the business criticality of that metric.

The contract does not include any formal provisions for the verification of performance reports submitted by Origin. With very few exceptions, users observe the attainment of service level requirements in the course of their normal work, even though they do not take proper measurements. Therefore, painstaking and costly verification by ICI is not needed. Having said that, the contract specifies regular audits (at least yearly) to be carried out for the benefit of ICI. Moreover, ICI has the right to initiate limited audit investigations upon suspicion of contract breach or other problem in the delivery of services.

2.3.2. Incentives for Service Level Attainment

2.3.2.1. Implicit incentive patterns

The audit procedures mentioned above, in combination with the possibility of benchmarking explicitly and in detail provided for in the contract, enable ICI to restore information parity whenever necessary and thus mitigate potential adverse incentives that might arise as a result of private information on the part of Origin. In addition, the contract requires ICI to provide Origin with assistance, access to information and to staff, to train its own staff to the proper use of the systems and to generally support Origin in fulfilling its contractual obligations. This clause serves as a deterrent of opportunism (moral hazard) from the part of ICI (see chapter 5, section 4.1.3 and chapter 2, section 4, paragraph 4).

Origin's liability is mainly defined in terms of financial penalties payable to ICI, depending on the services affected and the extent of under-performance or stoppage. Origin's other liabilities are each limited to 20% of the total value of software products supported by Origin that year.

Most computing assets were leased by ICI and these leases have been transferred to Origin. In addition, 400 staff have been transferred under the TUPE (Transfer of Undertaking - Protection of Employment) regulations. Infrastructure assets (mainly buildings) have been leased to Origin for ten years. The contract includes extensive provision for the allocation of rights to the software. In short, the ownership of intellectual property rights for software (owned by either party or developed, bought or licensed from any third party) is not altered. Instead, licences to use and modify software are transferred or granted to Origin and ICI as necessary. An exception to this is software (or enhancements) that Origin may develop for the sole benefit of ICI during the course of the contract. Intellectual property rights (ownership) for such software are assigned to ICI and ICI, in turn, licences it to Origin as required for the provision of services. Upon contract termination, licences are recalled or adjusted as necessary and ICI seeks appropriate licences from Origin to use whole or parts of software that relates to ICI services. This software includes system and application software as well as system development tools and methodologies.

In other words, the contractual provisions relating to the ownership of software do not seem to address any incentive implications of software ownership, other than practical matters pertaining to usage for service provision, while preserving the original property rights. To be sure, the contract refers to software or enhancements that Origin may implement in order to improve its management of services and considers the possibility of ICI investing in such software in anticipation of some agreed benefit to ICI. However, this is seen as an exceptional possibility, rather than an obligation or expectation. There are at least two conceivable reasons why incentive implications of asset ownership are ignored. First, in this case no revenue stream accrues to software ownership *per se*. Origin's income is defined separately in the contract and most incentives depend on the structure of compensation. Second, Origin's obligations are extensively defined elsewhere in the contract. Therefore, is less need to induce performance implicitly by means of asset ownership.

On contract termination, ICI retains the right to encourage those Origin employees engaged exclusively in services to ICI to take up employment with ICI. Moreover, the contract explicitly licenses Origin to use ICI-specific know-how. Such know-how and additional knowledge that Origin develops in the course of providing services to ICI reverts to ICI upon contract termination. Confidentiality clauses are in place to protect information, documentation and intellectual property rights on exclusive or restricted software from misuse by either party.

According to the Group IT Manager, the greatest incentive faced by Origin is the prospect of more contracts with ICI businesses. Since the data centre contract is due to expire without renewal, given the opportunity to bid for new system developments, Origin values the future relationship with ICI and strives to build and maintain a good reputation.

2.3.2.2. Contractual reward/penalty devices

Origin's compensation is based on a fixed price with appropriate price escalators (market indices) to take into account changing market prices. This is a powerful incentive, typical for standard services, which is also combined with ICI's emphasis on market benchmarking aimed at ensuring competitive pricing.

Having said that, charges are specified on a time and materials basis but for a set volume of computing and human resources. Moreover, payments to Origin are bounded, both at the top and bottom ends. In particular, the contract forecasts the volume of computing power for each of the five years of the contract in decreasing amounts (since the mainframes are being phased out). ICI is required to review demand forecasts on a quarterly basis for a rolling twelve-month period. A fixed value of total additional services (including projects that already were under way) was also agreed at the time of contracting, for the whole period of the contract. Unit prices were also agreed for adjusting payments to actual volume variations.

The contract guarantees a minimum threshold of payments equivalent to 90% of the forecasted volume of basic services for each year, as well as a minimum threshold of payments for the total value of additional services throughout the duration of the contract. This threshold for additional services can be reduced by one third of any excess revenues generated on basic services over the course of the contract. In the case of divestment of any ICI business unit the relevant thresholds are reduced proportionately, unless Origin does not continue to provide the services to the divested entity, in which case Origin is guaranteed 75% of the divested proportion of the value of the services. If actual demand exceeds the estimated volumes, any excess of up to 20% of the estimated value of basic services can be apportioned and deducted from the minimum threshold of subsequent years. In the case of excess actual demand, additional discounts are applied, on the basis of scaled proportions. Penalties for poor performance are excluded from all these calculations.

A recent contract amendment revokes the minimum threshold for additional services, the discounts and deductions in cases of excess demand and the guaranteed revenues in case of divestment. Only the minimum thresholds for basic services are sustained. These amendments are associated with extensive corporate restructuring that has been taking place in ICI over recent years.

Two kinds of explicit performance incentives are defined in the contract. First, a penalty system deters Origin from falling short on the set performance targets ('performance index rebates'). The mechanism specifies escalating percentage charges on the business value of the software, depending on the type and severity of the each performance shortfall. Fourteen metrics on the quality of service delivered, are measured with respect to each of the software products. Two additional metrics relate to contract management. For each metric, defined ranges of actual performance are classified as 'marginally deficient', 'deficient' and 'seriously deficient', depending on the respective targets. A percentage is assigned to each performance category, namely 3%, 7.5% and 20% respectively. Each metric is also assigned a relative weighting. The penalty for poor performance in any

metric for any software product is calculated as the product of the deemed value of the software product to the business for the given review month (if the metric is reviewed monthly), the weighting of the metric and the performance category percentage.

Second, Origin is liable to pay penalties for outages (unavailability) with respect to any application. An outage is defined to commence when response time rises higher than a set level (including total unavailability). Penalties ('liquidated damages') are payable when an outage or a series of outages last more than a defined period of time, during defined working hours. Charges are defined for each software product in an increasing scale for each day of outage. A premium charge is applied with respect to some applications for outages occurring at the end of a month (when demand and criticality are greater).

Performance index rebates are limited for any one month to 2.5% of the total value of the software products for the given year. In addition, Origin's total yearly liability to pay any penalties is limited to 14% of the same value. The application of the two kinds of penalties is exclusive (i.e. not additive) with respect to any one application or software product and the penalty of highest value is applicable. Whenever any penalty is payable, Origin has to take action to correct the problem and to prevent it from recurring. If penalties become payable repeatedly within defined rolling periods then ICI can treat this occurrence as consistent failure and examine the option of terminating the whole or part (i.e. certain services) of the contract on grounds of breach.

All explicit incentives are deterrent (penalising), as opposed to encouraging (rewarding) incentives. As mentioned in section 2.3.1 above, there are some service level requirements involving considerable uncertainty, which would be suitable for encouragement-type incentives. However, if such service level requirements are few and exceptional the impact of the incentive on Origin's total compensation would be negligible. Therefore its effectiveness would be constrained and it may not have been worth the contracting and administrative costs.

2.3.2.3. Other contractual and organisational safeguards

According to the Group IT Manager, ICI is satisfied both with the operation of the penalty mechanism and the performance of Origin, as no serious problems have occurred. The most important incentive for Origin is its reputation within ICI, which arises from the possibility of Origin bidding for and winning contracts for new system developments throughout the group. Both parties value and rely on the personal relations between contract managers for the smooth operation of services and the successful fulfilment of the contract.

ICI considers the change management procedure as effective in controlling opportunistic charging of work as out of the scope of the contract. This is further evidence of the relationality of the contract, which attempts to curb opportunism through *ex post* adaptive processes rather that *ex ante* incentives.

Finally, the contract defines in detail the conditions under which it can be terminated by either party. Extensive terms are also included for managing the transition to a new service provision regime, in case of termination of part or whole of the contract. The conditions of termination include, among others, the occurrence of a consistent failure. However, consistent failures are extreme situations. Besides, ICI has much to lose from a premature termination due to consistent failure. Therefore, the incentive effect of this clause is marginal.

2.3.3. Organisation and Management of Service Provision

2.3.3.1. Roles and responsibilities

The contract defines in detail structures and processes for the management of the contract and the services. For this purpose, both parties commit a management team. Each team's and each individual's responsibilities are also specified.

ICI's management team comprises 2.5 (full-time equivalent) contract managers at the corporate level and another contract manager from each business unit. In addition, each business unit assigns a 'system owner' occupied with overseeing the services associated with each software product used, from a user's point of view.

Origin's management team comprises the relationship manager, the corporate account manager, the corporate contract manager, the contract managers assigned to ICI business units and the implementation manager. The relationship manager has overall responsibility for the relationship between the parties and the other members of Origin's team report to him. The corporate account manager is responsible for available services and for arranging proposals to meet ICI's additional requirements, according to the contract change management procedures. Corporate and business unit contract managers are responsible for ensuring Origin's compliance with the contract. The implementation manager was responsible for the necessary activities during changeover when the contract came into force.

2.3.3.2. Coordination and communication mechanisms

The contract also defines four regular meetings. A high level management meeting is aimed at reviewing the overall progress of the relationship on a quarterly basis. All members of both management teams are normally required to participate. Corporate and business unit service review meetings are held monthly to review compliance with the contract, service level requirements and plans for the next period. All corporate and business unit contract managers attend the corporate service review meetings whereas only the relevant contract managers attend business unit service review meetings. Product review meetings by business unit are held monthly or as needed to review issues specific to a particular software product. The relevant business unit contract managers and system owners attend these meetings.

Monthly performance is managed 'by exception' to minimise the effort spent on monitoring. In other words, performance reviews are rather trivial processes, focusing on any problems that may occur, rather than painstakingly examining every aspect of performance. This requires, first, that service level requirements be fairly standardised and stable, second, that the performance measurement system is considered sufficiently efficacious third, that recurrent or persistent problems are not common and fourth, performance is not misrepresented. The inherently standard and predictable nature of data centre operations facilitates this economising approach to performance monitoring.

2.3.3.3. Locus and scope of decision making

Another important decision making process is the change management procedure which sets out in detail how changes to the contract or the services can be initiated, negotiated, agreed and implemented (as described in section 2.3.1 above). System owners can initiate some enhancements but the procedures are mainly carried out by business unit contract managers and controlled by Origin's corporate account manager and ICI's corporate contract manager.

An appeals procedure (escalation process) has been instituted that refers serious problems in service levels up to the level of the ICI Group IT Manager and Origin's Relationship Manager, who use their authority and direction to restore service provision. A dispute resolution procedure is separately set up to handle disagreements of any kind (e.g. with respect to applicable charges or the actual service performance). Such disagreements are in the first instance raised at a regular or extraordinary management meeting. If no agreement is reached, the matter is referred to senior business managers from both parties (e.g. finance directors) who are asked to "*cooperate in good faith to resolve the dispute as amicably as possible*". The last resort before litigation in court is to engage third party mediators and/or experts, depending on the nature of the dispute.

Clearly, the contract makes every effort to prevent costly litigation which is also likely to have implications for service quality. Again, this is evidence of the relationality of the contract. Nevertheless, the convivial personal relations among the management teams of both sides are considered to be a more powerful means of resolving problems. This is a reference to the power of compatible corporate cultures (as defined in chapter 2, section 7 and included in S-CAGE in chapter 5, sections 4.2.1, 4.2.3 and 4.3.2) to address contingencies that arise as a result of contractual incompleteness.

During the first year into the contract, the focus of the contract management team's work at the corporate level has been to manage the transition from the in-house operation to the outsourced organisation and eliminate any problems. During the second year, emphasis shifted to sustaining competitive supply, mainly through market benchmarking. Half-way into the contract, a review has been initiated at the corporate level with a view to adjusting existing arrangements and strengthening the 'partnership' with Origin. A shifting emphasis of management attention over time is, therefore, apparent.

2.4. Comments

The view from the top as presented by the Group IT Manager appears to lend credence to the neoclassical economic argument that the division of labour grows with the size of the market. ICI appear to be making their strategic IT procurement decisions on the basis of market capabilities and competitive conditions. These can be seen as evolving over time with the growth of the IT market and the standardisation of technologies and systems.

Issues of asset specificity enter into the relevant decisions indirectly. Every organisation, in the process of demarcating their strategic core processes and capabilities implicitly define what is unique (i.e. specific) and what is not. Moreover, the capabilities offered by the supply market are in part (or at least at equilibrium) determined by the size of the respective demand. The unavailability of

a competitive market for a given service can be an approximation for the specificity of that service (assuming that the market does not face any other distortions). In any case, whether asset specificity and related microanalytic considerations have played a role in ICI's strategic decisions to pursue outsourcing or not (a binary decision), they have been pertinent in structuring the contracts.

The case of ICI suggests that neoclassical economics and microanalytic institutional alternatives can be usefully distinguished in terms of the level of analysis to which they refer. Inasmuch as strategic decisions between in-house provision and external procurement are based on an aggregate and more long term perspective, neoclassical microeconomics may provide sufficient explanations. Accordingly, the microanalytic stance of the S-CAGE framework turns out to be more relevant to the exact structuring of governance, given the strategic direction. Findings presented here highlight the bottom-up orientation of the S-CAGE framework which makes it pertinent to the study of contractual provisions. However, this is not to say that the latter considerations do not inform and do not have an impact on the strategic decision. This argument makes an important qualification to the theoretical basis developed in the conclusion of chapter 2. In other words, whereas there may be microanalytic imperatives for organising IT provision in one way or another, the strategic decision depends on an array of other parameters as well.

By isolating data centre operations from any other services, ICI has managed to set up a relatively more discrete contract as compared to total outsourcing arrangements. The initial competitive bidding process has led to a fixed price contract that maximises incentive intensity. The latter is further strengthened by commitment to market benchmarking. This has been possible due to the comparatively standard nature of data centre operations that can be matched against comparable operations in the market. Despite the costs of switching suppliers and the resultant advantage that Origin has as the incumbent supplier[3],

[3] The extent of the comparative power of Origin as the incumbent supplier and the corresponding costs of switching to an alternative supplier is a measure of the investments involved. The

asset specificity is relatively low. Moreover, the structure of the penalty system and Origin's limited liability show that service level requirements are fairly straightforward, involving relatively little uncertainty. There are too few uncertain, company-specific or innovative requirements to justify an encouragement-type (reward-based) incentive (see section 4.2.2 in chapter 5).

The opinion of the Group IT Manager also highlights the importance and operation of reputation effects (see chapter 2, section 7 and chapter 5, section 4.2.1). Origin values the future relationship with ICI as they have the opportunity to win additional contracts, given that current operations will be discontinued. To that extent, ICI can trust that Origin will continue delivering high quality of service. The calculative nature of trust is, therefore, manifest. The role of culture is also highlighted by the importance attached to the personal relations among contract managers of both parties in terms of resolving conflicts (see section 4.3.2 in chapter 5).

Overall, it is evident that most of the questions asked in the interview guide were answered as expected by the S-CAGE framework. A few important unanticipated results are discussed in the conclusion to this chapter (section 4). The contract provides a wealth of information, albeit it is structured in a legalistic way that does not lend itself to immediate interpretation from the point of view of S-CAGE. A significant amount of investigation and cross-examination of the different schedules and clauses included in the text is required in order to arrive at informed answers to the questions posed in the interview guide. However, the result is rewarding, as it reveals mechanisms, connections and interpretations that are deeply hidden under the jargon and the exhaustive stipulations.

---

transaction costs of negotiating another contract and installing another supplier can also be thought of as a consequence of the specificity of data centre operations to the respective business processes.

3. Total IT Outsourcing at Anglian Water

Anglian Water Plc provides water and sewage treatment services in the UK and abroad. Anglian Water Services Limited is the principal subsidiary, regulated by the 1991 Water Industry Act. It is geographically the largest of the ten regional water services companies in England and Wales, serving a population of over 4 million in the east of the country. The Group's turnover for 1997 was £837 million, nearly £700m of which was generated by Anglian Water Services Ltd. Anglian Water International Limited is the other main subsidiary with operations predominantly in the American continent and the Far East.

For the purposes of this research, contact has been made with the Service Manager of Information Services of Anglian Water Services Ltd (henceforth referred to as AW). This is the key person leading the management of a major IT outsourcing contract with CSC.

3.1. Historical background

Until recently, AW's IS/IT infrastructure was fragmented with many standards and systems for essentially the same processes across the organisation. They were operating a Bull mainframe for administration (payroll, accounts, billing etc.) and a VAX mainframe for production operations. As a consequence, there were effectively two large IT departments within the business with variably overlapping responsibilities.

The present IS Director arrived at AW shortly before October 1994 and made a substantial contribution to the direction of the IT review and to the decisions that ensued. AW runs a business planning process from October to February every year. In October 1994 they initiated a review of existing IT infrastructure and decided to migrate to client-server technology while also rationalising their IT operations. The planned changes for IT were also accompanied by a broader restructuring exercise

and cultural change throughout the organisation. This organisational change involved, among others, empowering and training programmes for all staff.

CSC was one of the consulting firms that were invited to contribute to the business and IT planning processes. By February 1995, AW had decided to outsource all existing operations and new developments, apart from telemetry (see below), to CSC. They immediately entered contract negotiations. The Service Manager got involved in the process in August 1995, coming from Anglian Water International Ltd. Her main task has been to lead the development of service level agreements and the management of the contract. By that time, the framework agreement was completed, specifying a 10 year contract, the first 5 years of which would be spent on developing new client server applications while managing the existing legacy systems, as well as the transition. This is a major project for AW, involving complete replacement of all existing IT systems, starting from a clean slate. Over the second half of the contract, CSC is expected to manage the operation of the new systems. Contract negotiations were complete and the contract was signed on 31 October 1995.

3.2. *Governance Portfolio*

The in-house team manages four main tasks: the contract with CSC, the telemetry services, IS strategy, data protection and security. They may seek the advice of CSC regarding some or all of these matters but it remains AW's responsibility to control and deliver them.

The telemetry systems involve communication with and control of various automatic devices (pumps, sensors etc.) at about 8000 unmanned sites (wells, reservoirs etc.) all over the region covered by AW. The telemetry network uses TCP/IP over low power radio and microwave links transmitting information to and from a central control room. A large proportion of the 60 people running the telemetry services are engineers spending most of their time in the field looking after the equipment. There is a technical interface between the telemetry systems

and CSC's area of responsibility. This seems to be a well defined boundary, although they sometimes have to manage the organisational coordination. AW consider their telemetry systems as state-of-the-art and do not believe they have anything to gain from employing contractors to manage them. Therefore, their operation has been kept in-house. Moreover, such systems are highly specific to the fairly unique nature of AW's operations. Internal organisation is thus justified on transaction cost grounds, while it is also unlikely that a large, competitive and more efficient market can be sustained given the specific nature of these systems.

3.3. CSC contract: Governance structure

3.3.1. Assurance of Service Delivery

AW's contract with CSC serves about 170 sites and covers virtually all IT services, including networks, mainframe operation, desktop support, help desk and new system development. Services and Service Level Agreements (SLA) are specified entirely in business terms. Achieving this had been a difficult and lengthy process. The management of IT prior to outsourcing was not based on any formal performance requirements. These were developed from scratch and particular effort was made to ensure early user involvement in specifying them.

The emphasis on business-based performance measures is also suggested by the term 'Key Business Service Components' (KBSC). There are about 120 KBSCs representing elements of service being measured. The derivation of the SLAs associated with new developments becomes part of the whole design and development process of each project. This approach emphasises that users understand what to expect from CSC and how/why they earn penalty points (see below) when CSC does not meet performance targets. This has also practical implications, as it is usually the users who first raise a complaint about a problem whenever they face it. They should therefore be able to understand when to flag a complaint. Overall, it is primarily a user-driven process and CSC's help desk plays

a central role in recording relevant information about calls and responses. This also means that if users are not inconvenienced CSC will not be penalised.

CSC is responsible for measuring and reporting performance on these KBSCs, mainly via the help desk. They are also responsible for calculating and reporting penalty points. CSC also takes action to verify this information with the users through user surveys that the help desk administers. This user survey is not part of the obligations of CSC, it is above and beyond the contract. The service management team is responsible, among others, for policing this process. They hold weekly meetings with CSC account management staff to monitor performance issues, penalty points and persistent problems.

3.3.2. Incentives for Service Level Attainment

AW's legal advisers interpreted the legislation covering the licences of water utilities to include IT assets, namely that these assets cannot be transferred to a third party. Therefore, hardware was not transferred to CSC. Other reasons for not transferring the assets relate to accounting and financial issues, mainly that the fixed asset base of water utilities contributes positively to the pricing formula established by the regulating authority. Therefore, AW had an interest in maintaining ownership of these assets, although they constitute a relatively small percentage of their total asset base. Although residual ownership (see chapter 2, section 5) remains with AW, the contract attempts to restore CSC's incentives, which otherwise accrue naturally to ownership.

According to the Service Manager, CSC is encouraged to manage these assets as if they owned them, in order to take advantage of potential efficiency gains. For example, AW does not in principle object to possible decisions by CSC to move or share the data centres, but they reserve the right to approve such plans. Sovereign House at Histon, Cambridgeshire, is a CSC site but also houses AW's Information Services team of 15. Other sites, for example a Data Centre in Lincoln is owned by AW but now managed by CSC, more or less with the same (transferred) staff. On

the whole, 200 staff were transferred to CSC, of which 120 were permanent and 80 contract staff. Software licences were also transferred.

CSC's compensation has a fixed cost component (subject to a cost inflation index) which is reduced year after year, plus a variable element associated with new systems development work. This compensation is subject to a profit sharing clause whereby any improvements that CSC achieves above a fixed percent profit margin, are equally shared with AW. Although this compensation arrangement is of the cost-plus kind, it imbues an incentive due to the fact that costs are agreed in advance. Because compensation for new developments is fixed in the contract regardless of the progress made by CSC, the sooner CSC delivers the new systems, the more they benefit and vice versa.

Penalty points are awarded on an exponential scale, depending on the criticality of the service. Criticality appears, in general, to be a function of the effect of a problem to important business operations and the public image of AW (e.g. billing), which can be a source of significant damage. When a problem occurs, the prescribed penalty points are awarded. Often, the initial penalty is nil, particularly when users are not affected (e.g. when a fault occurs during the night). Delays in fixing the problem trigger exponential growth to the penalty points whereby the initial penalty points are multiplied by 2^n, where n ranges from 0 (at first occurrence) to a maximum of 5. Penalty points are periodically translated to monetary charges.

AW's contract provides for an open book policy and gives them audit and benchmarking rights. Moreover, AW wishes to encourage CSC to innovate and to generate improvements. Having said that, it is not entirely clear to them how to implement such an incentive, both in contract clauses and in the overall service delivery process. AW and CSC have agreed in principle on the desirability of a reward-based incentive system (in addition to the current penalty-based incentive). CSC was to bring forward a draft recommendation, but it had not been submitted by the time of data collection.

3.3.3. Organisation and Management of Service Provision

At the time the research took place, AW's IS Services group had sixteen members. The IS Director reports to the Finance Director. The IS Strategy Manager is mainly monitoring the IT market, although it is not his responsibility to make technical decisions (this is CSC's task). Two members of staff provide general administrative support to the Information Services group. Figure 6.1 depicts an outline of the group's structure.

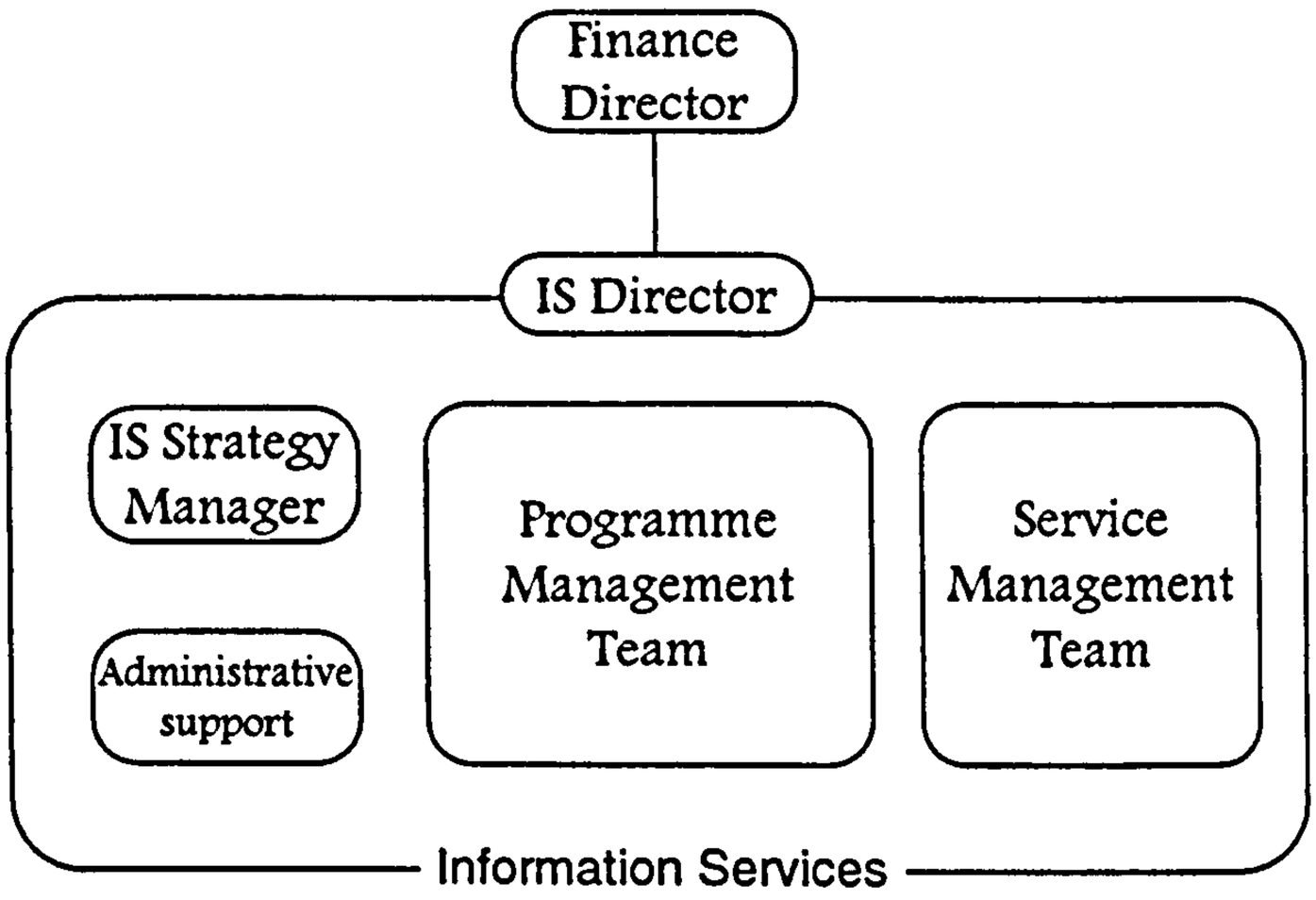


Figure 6.1

The Organisation of Information Services at Anglian Water Ltd.

There are eight members of the Programme Management Team. These are the key people facilitating the new client server development projects, liaising with and intermediating between CSC and AW businesses. Their responsibility is not to actually manage the projects (this is the responsibility of project teams comprising business representatives and CSC staff). Instead, they manage and integrate the process, ensuring that both users and CSC commit adequate effort, time and resources to the development process. Part of the Programme Management Team's task is to ensure that CSC's proposals make a good business case, are understood by the users and are revised according to user feedback. Moreover, they negotiate

with CSC over specifications and costs and sign invoices for payments once work has been delivered.

The Service Manager is heading a team of two service analysts and administrative support personnel looking after and policing the implementation of the CSC contract. The derivation and negotiation of the KBSCs and of the penalty system with CSC in collaboration with users has been a major exercise for the Service Management Team. Once KBSCs are in place, their attention is directed towards benchmarking (by employing specialist consultants), testing and perhaps renegotiating them. They see the whole outsourcing project as a learning experience and they foresee that over time they will be emphasising different aspects of the relationship management process. The Service Management Team also liaises with the legal department to receive legal advice and support on issues pertaining to the contract. Finally, they administer the internal mechanism for charging AW business units for their use of CSC services. This mechanism also incorporates adjustments as a result of penalties incurred by CSC.

Most members of the IS group are relatively new to AW, notably the IS Director and the Service Manager. Programme Managers come from diverse backgrounds such as engineering and finance. Only the IS Strategy Manager and two of the Programme Managers came from the pre-outsourcing IT organisation.

One of the main tasks of the programme management and the service management teams is to ensure that proposals from CSC make business sense, represent value for money, are in line with the agreed IS strategy and are understandable to the users. In this respect they do not seem to be concerned with the fact that many of their technical skills have been lost to CSC, as long as they can manage the process in business terms. For AW, technical expertise is readily accessible to external sources. The new developments are process oriented and cut across the vertical internal boundaries between business units and functions. This has implications for the coordination of systems and business units.

It appears that AW and CSC often engage in trivial renegotiations or minor disputes at a micro level, while at a macro level AW wants to ensure that overall service levels and costs are satisfactory. However, the contract specifies an escalation process starting from the Service Manager, to the IS Director and up to the Finance Director and their CSC counterparts. While there are points of friction in the relationship with CSC, AW emphasises the importance of managing a sustainable relationship.

3.4. Comments

An interesting feature of the AW case is the fact that they opted for total outsourcing. Although total outsourcing has been common, especially during the early 1990s, it is worth exploring the possibility of an explanation within the S-CAGE framework. The reasons for total outsourcing are not readily apparent, although the early involvement of CSC in the strategic decision making process must have played a role. Other stated reasons include the lack of in-house skills combined with insufficient in-house credibility to deliver the massive redeployment of IT resources that had been decided. Moreover, they saw the outsourcing option as an opportunity to buy in best practice and to re-engineer the IS department. Although these reasons seem to justify outsourcing, they do not explain why AW assigned all IT activities to one supplier instead of employing different vendors for different services.

What they really outsourced was not the operation of a fixed set of services but the replacement of practically all existing systems. The alternative to total outsourcing could possibly involve employing a number of contractors working on different development projects, another set of contractors managing existing operations and a third set of contractors to manage the operation of the new systems. AW would then have to keep all the technical decisions in house and coordinate both the technical as well as the commercial aspects of the work of all those third parties. The task of coordination could have proven to be overwhelming, with the only feasible alternative being total (or nearly total) insourcing of the skills required for

the new developments. This coordination burden combined with the perceived lack of skills[4] and the desire to reorganise the IS department, can provide a partial justification for total outsourcing.

This is an extreme case where the only feasible clustering strategy appears to be one cluster for all services. It seems to be a result of the heavy interdependence between existing operations, new developments, the transition and the new operations. New developments are not isolated projects as they are all based on the same technical platform and cut across internal organisational boundaries. Substantial business process redesign across the organisation is also involved. Because there are skills and decisions that are common to all development projects and specific to AW, it makes sense to assign responsibility for them to the same party[5].

Also, given that the quality of service in development and in operation/maintenance are positively correlated (the respective costs are negatively correlated), it is also reasonable for incentive purposes to assign both activities to the same vendor[6]. In this way, the developer (CSC) is motivated to perform better so as to reduce the costs of operation and maintenance. Even more so, if the contract for operation is on a fixed price. Having said that, the Service Manager anticipates that perhaps halfway into the contract (when development projects are expected to be completed) further negotiations may be needed to adjust the agreement on systems operation. This also reflects an awareness of the need for adaptations as circumstances change over time.

Although the above is a plausible explanation for assigning all development and operation to the same vendor, it is not clear why the operation of the mainframes being phased out should also be the responsibility of the same supplier. This

---

[4] See the discussion on the firm as a knowledge repository in chapter 2, section 6.

[5] See the theory of property rights in chapter 2, section 5.

[6] See chapter 2, section 4.1 paragraph 9 and chapter 5, section 3, paragraph numbered 5.

arrangement assures CSC's revenue, despite the phasing out of legacy systems. On one hand, this assurance provides additional incentives to CSC to make AW-specific investments for the long term. On the other hand, it dampens its incentive to perform well in operating the mainframes. These relationships are quite subtle and they depend on the material conditions affecting each task and on the exact terms of the agreement(s). Moreover, there may be other, non-contractual, factors involved in these decisions, such as the strategic intent of each company, their relative bargaining positions and inter-firm politics. Such factors are not directly addressed by the S-CAGE framework.

According to the theory of property rights, the fact that AW has continued to own its hardware assets may create distortions to the incentives of CSC. At least, it definitely creates a basis for possible haggling over decisions to move, share, alter or sell them as CSC's capacity requirements change over time. Incentives which otherwise accrue naturally to asset ownership are, in this case, restored by contractual agreement. Given that the source of incentives is in residual (i.e. non-contractible) control rights, this is achieved by specifying joint decision making procedures to address instances where *a priori* rights are insufficient to cope. In practice, the contract gives open-ended rights to CSC to control computing assets (contractual agreement), subject to AW's approval (*ex post* decision making procedure). However, the incentive implications of asset ownership were not explicitly recognised.

The penalty system currently covers existing operations and its aim is to safeguard a minimum level of service quality and user satisfaction. The emphasis on business oriented service level requirements is noteworthy in this respect. Since the new system development projects comprise the largest part of the contract, the latter is of the cost-plus kind which is typical of requirements contracting for systems development. Similarly, the many uncertainties of these projects do not allow the introduction of formal incentives. Instead, the programme management team plays a key role in managing the process, including contingencies, in order to safeguard quality and cost.

4. Conclusions

The aim of this chapter is to demonstrate the use of the S-CAGE framework in analysing and understanding the governance of IT provision. The exposition of the two cases of ICI and AW have shown how S-CAGE can shed light on the practices observed in each firm and the differences between the approaches employed by them. Moreover, it has been shown how the main parameters of such contracts can be highlighted in an organised manner using the classification provided by the three layers of governance.

The literature on IT outsourcing is almost unanimous on the importance of crafting appropriate contractual arrangements (see chapter 1, section 7). However, there is relatively little insight on what makes a contract appropriate. Various attempts to systematise a classification of sourcing alternatives have demonstrated the availability of many options but have not produced a comprehensive taxonomy, beyond 'lessons learned' and 'best practices'[7].

A fundamental argument made in this thesis is that the quest for such a taxonomy is futile on account of the multiplicity and dynamic nature of the contingencies determining the IT sourcing strategy. Not only are there numerous strategic, financial, political, economic and technical factors, but the nature and relative importance of these factors changes over time and across, countries, industries and organisations.

---

[7] Some IT sourcing possibilities are quoted in chapter 1, section 2. Some others are reviewed by Willcocks and Lacity (1998: 1-41): contract-out, preferred contractor, buy-in, preferred supplier, alliance, insourcing, offshore outsourcing, IS department as vendor, partnership, selective sourcing, transitional outsourcing, total outsourcing, value-added outsourcing, equity holdings, multi-sourcing, co-sourcing, spin-offs, creative contracting. These terms are drawn from lessons learned in practice and/or from what has been advocated as industry best practice. The definitions of these terms are not always unambiguous and, consequently, they have little discriminatory power for theoretical purposes. For the same reason, they give limited guidance to practitioners.

Instead of seeking a catalogue of sourcing alternatives, the S-CAGE framework presents a classification of the elements of contract. Whatever sourcing strategy an organisation adopts, the structure of the contract can make the difference. The value of the framework lies in its capacity to expose, analyse and question the provisions made in an IT contract, in a way that is relevant to management and not just to the legal professionals.

A key feature of the S-CAGE framework is the notion of clustering. This is not a totally new idea. It has similarities with the notions of selective and multiple sourcing. Also, companies have always been bundling IT services, applications and projects in various ways and procuring them from different suppliers. The argument for service clustering presented in chapter 5 is elaborated in terms of the theoretical basis developed in chapter 2, section 8.1. In particular, the criteria for clustering or separating services draw on microanalytic economic/contractual issues, such as asset specificity, incentives and measurement problems. However, the case studies in ICI and AW have revealed some explanatory limitations of this approach.

Although ICI and AW have adopted outsourcing for the transition from mainframe to client-server systems, they have followed a different clustering approach. The reason explicitly quoted by ICI is that they wanted to offer stronger incentives to Origin, in the form the prospect of additional contracts. AW saw the whole exercise as a single project from the outset. As discussed in previous sections (mainly 2.4 and 3.4), the role of incentives, coordination costs, skills/knowledge and asset specificity provide partial explanations for the clustering approach followed each case.

The differences in the incentive effects of the clustering approaches adopted by AW and ICI are also worth exploring. By soliciting separate bids for each new project, ICI aimed to generate strong reputation effects as an incentive for Origin to perform well during the operation of the legacy systems. In contrast, AW assigned new developments to CSC from the outset. In comparative terms the incentives faced by CSC with respect to mainframe operations in the later case are weaker.

The case studies indicate that other factors, outside the scope of the S-CAGE framework, impinge on the clustering of IT services for the purposes of outsourcing. For example, an obvious explanation for the different clustering approach in ICI and AW is the size and organisational structure of each. ICI's contract was entered into at the corporate level and all new systems development projects are being undertaken independently by individual business units. AW is a much smaller company. The contract was entered into at the business unit level (Anglian Water Services Limited) and affects the whole business unit. Another factor is the strategic intent of senior management. For example, the Group IT Manager at ICI emphasised the importance of seeking market opportunities for good procurement agreements. Traces of political and historical antecedents to AW's decision can be found in the involvement of CSC since the early strategy review, which concluded with the decision to outsource.

Therefore, insofar as service clustering is concerned, the S-CAGE framework provides relevant and valid explanations (and directions) but it does not cover all the determinants. Having said that about clustering, the layers of governance framework proved to be very effective in analysing the contractual practices in each case.

The case study evidence shows that the condition of service specificity can manifest itself in many different guises within a contract. This testifies to the elusive nature of the concept and, thus, to the difficulty of measuring it in quantitative studies. Various approximations for asset specificity have been proposed in the literature. These range from multidimensional instruments asking questions on the character of the particular transaction to secondary sources such as the size and competitiveness of the supply market. It emerges from these case studies that a good measure (or at least indication) of specificity is the contractual provisions that exchange parties introduce in order to safeguard their specific investments. Part of the elusiveness of specificity is its subjectivity (i.e. what is specific for one organisation may not be specific for another). Therefore, a rational contracting party would make every effort to protect its specific investments. The various

clauses in ICI's contract reviewed in section 2.3.1.1 above are evidence to this. The reliability of contractual clauses as a measure of specificity is limited by the parties' opportunistic attempts to negotiate more in their own favour and their limitations to rationality.

In both case studies, emphasis is placed on specifying appropriate performance measures and targets while minimising the costs of administering them (measurement and information costs). The effort made by AW to concentrate on business-oriented measures (Key Business Service Components), is worth noting. Their intention in doing so was twofold. First, to distance themselves from technical decisions as much as practically possible (see the relevant discussion in chapter 5, section 4.1.1). Second, to enable end-users to understand what to expect from their IT supplier. Apart from ensuring commitment, managing expectations and, hence, improving the likelyhood of perceived success, this action facilitates information parity, performance monitoring and contract management with minimal additional administrative procedures.

Both companies employ incentive mechanisms that penalise the vendor in cases of performance shortfall. An implicit reward in the AW-CSC contract derives from the fact that costs for new systems development projects were agreed and fixed in advance. Therefore, an incentive is derived, as if the contract was on a fixed price. Inasmuch as CSC cannot extract additional compensation for cost overruns, it has an incentive to deliver systems faster. Acceptance procedures must then safeguard system quality. In the case of ICI, the opportunity for rewarding incentives was identified with certain service level requirements related to the accuracy of estimates and plans for new projects. These requirements are a contractual obligation but no financial penalty is attached, in recognition of the uncertainty that characterises them. An incentive providing additional compensation when Origin overcomes the uncertainties, would be suitable in this case. However, these requirements are a very small part of expected performance (although potentially significant in terms of value) and any related incentive would not make a big enough impact on Origin's total compensation to justify its introduction.

Although in both cases appropriate provisions are made specifying the conditions under which contract termination can be invoked, the threat of termination is not seen as an important incentive. The conditions for termination specify a very low performance threshold and both ICI and AW prefer not to get close to that level, since it would involve significant damage to their businesses. Actually invoking termination on grounds of poor performance would be even more damaging to both vendors and clients. The fact that both parties prefer to avoid termination, dampens the client's resolve to enforce it and consequently reduces the incentive effect of the termination clause. This suggests that the more specific investments are involved and the more relational the contract is, the less powerful is termination as an incentive (and *vice versa*).

Similarly, the incentive importance of asset ownership was disregarded both by the interviewees and by the contracts in both case studies. Any transfers of property rights seem to have purposes not related to incentives. As it turns out, companies give priority to two other issues over incentives. First, the contracts protect the intellectual property rights on software. Second, any asset transfers are made to protect the continuity and operation of services. Once these two purposes are taken care of first, there remains very little scope for reassigning ownership rights for incentive purposes. It appears that although the theory[8] provides an elegant and credible proposition, in practice other factors seem to take precedence. Of course, this observation is drawn from the particular kinds of contracts studied here. The incentive significance of asset ownership in IT contracts warrants further empirical investigation.

Finally, the relational character of these contracts is manifest in the numerous clauses that defer decisions, plans and actions to joint decision making procedures during contract execution. Often the contracts call on 'all reasonable endeavours' of the parties to take into consideration each other's interests and to resolve

---

[8] See chapter 2, section 4.2, paragraphs 2-5 and section 5, paragraphs 4-6. See also the relevant discussion in chapter 5, section 4.2.1.

disagreements in an 'amicable' way (see the the discussion on discretionary collaborative behaviours in chapter 5, section 4.3, paragraph 5). Moreover, the formal incentive mechanisms would not be efficacious without supplementary administrative procedures to control aspects of performance that cannot be contracted in advance. Such procedures are the change management and approval provisions found in both cases. These and other similar provisions are akin (even if qualitatively different) to the open-ended rights and duties afforded by the fiduciary regime of internal organisation (see chapter 2, section 2.2). In this way, markets replicate many of the adaptive advantages of in-house service provision, alongside penalty clauses and other incentive instruments (typical advantages of market structures).

The variations in the adaptive, incentive and other mechanisms instituted by the two contracts, which otherwise correspond to very similar services, provide empirical support for the concept of governance continuum developed in the end of chapter 2. In other words, two contracts that can, in general, be classified as being the same, may contain differences that have important managerial and economic implications. This is the essence of the microanalytic approach advocated in this thesis, namely to examine governance mechanisms in their details and to study how they may combine in various configurations yielding different governance structures along the governance continuum.

The two cases presented here demonstrate, albeit in a limited way, the capability of the framework to describe, analyse and interpret the governance of IT provision. As such, S-CAGE is also useful in understanding different practices between organisations. The implication for the relevance and applicability of S-CAGE is that it is relevant to a bottom-up analysis of the clustering approach and the details of governance from an economic/contractual point of view. Further research strategies for testing and enhancing the S-CAGE framework are explored among other conclusions in the following final chapter.

CHAPTER SEVEN

Concluding Remarks and The Way Ahead

1. Introduction

The widespread adoption of IT outsourcing since the late 1980s suddenly turned the spotlight on an aspect of IT management that had previously received little attention in both the academic and practitioner communities.Organisations have always been using contracts, of one kind or another, for systems and services but the issue was assumed to be somehow peripheral to the organisation and management of IT or to consist of more simple make or buy decisions. Given the unquestioned assumption that IT provision is the responsibility of large in-house departments or divisions, attention was centred around issues of centralisation versus decentralisation, end-user computing, and charge-back systems, to name a few. When some large corporations in the US started replacing all their IT operations with single contracts, many academics and managers were caught off-guard.

After consideration, some observers noted that this phenomenon was not so much a revolutionary change but rather an evolutionary step. Under the light of the theory developed in this thesis, the issues of centralisation versus decentralisation, end user computing and charge-back are simply elements of the governance of internal IT provision. They are about measuring performance, providing appropriate incentives and allocating decision authority and control. In other words, they are about adjusting arrangements in the three layers of governance, respectively, of the S-CAGE framework (Figure 5.4), within the general governance category of internal organisation. Therefore, the move from internal organisation to contracts of the total outsourcing kind is, analytically, a small shift on the governance continuum (Figures 2.2 and 2.3).

This is not to say, however, that the above is a complete account of the phenomenon on the whole or of individual decisions made by organisations. On the contrary, as chapter 6 concluded, the broader historical, technological, economic, strategic and political explanations and interpretations reviewed in chapter 1 are important complements to a comprehensive understanding of what has been going on. Similarly, various abstract or empirical exercises in the context of economic theories illuminate the potential workings of parts of the relevant economic and organisational system.

These inroads motivated the desire to fill a small gap in the middle ground between partial, restricted economic models and unifying, all-encompassing accounts. The focus is the governance structure. The perspective is provided by economics. The task is both abstractive and synthetic. The governance structure is abstracted from other co-determinants of the organisation of IT provision, in order to define its nature and examine the structure and workings of its constituent parts. In doing so, several relevant strands of economic theory are synthesised into an integrative whole, the governance structure. It was the task of chapter 2 to review the relevant strands of economic theory, to assess their implications for governance and to combine their often disparate means and ends into a concept of governance structure. Transaction cost economics, agency theory, the theory of property rights, theories of information, knowledge and measurement and economic treatments of culture, trust and reputation effects provided the raw material for this task.

The governance continuum, established at the end of chapter 2, is one of the two pillars on which subsequent theoretical propositions rest (the other being the clustering continuum). To be sure, both the selection of relevant theories and the notion of a continuum of governance emerged out of the action research at British Petroleum. However, it is felt that this separation and linear exposition aids a better understanding of the concepts involved and the arguments made. The argument made at the end of chapter 2 entails a definition of governance as a system of organisational and contractual arrangements, including performance

measurement, incentives and decision making and control. The composition and characteristics of the services being governed is singled out as the main determinant of the exact configuration of such a system, which, in turn, may lie anywhere between two fictional opposite extremes, the pure market and the pure hierarchy.

Chapter 3 stands as a pause in the flow of the thesis for self-reflection. It identifies the philosophical underpinnings of the research project in the realist philosophy of science and draws implications for the nature of the proposed theory and the methods employed. The upshot is that governance structure and its mechanisms are real entities existing independently of our knowledge of them and of their manifestation in the events that we observe. As such they possess causal powers that may lead to particular (observable) organisational and contractual arrangements, contingent on the actualisation of various unpredictable conditions under different circumstances.

For example, formal incentive mechanisms can take several forms: they may be rewarding or penalising, they may be based on nominal or aggregate performance and they may have lighter or heavier financial consequences, to name a few parameters. Their feasibility and effectiveness is conditioned by, among others, the nature of the respective services, other incentives facing the supplier and the willingness of each party to enforce the mechanism and bear the consequences. Therefore, whilst such mechanisms are potentially very powerful in inducing high quality and low cost of service, their exact role and configuration is indeterminate outside an actual context.

What is called for, therefore, is a comprehensive account of the potential capacities and ways of acting of the elements of IT governance which can take different forms under various conditions, leading to particular consequences. Such an account constitutes both a basis for explaining observed governance structures for IT as well as a guide for decision making in practice.

The exploratory study of the outsourcing practices and opinions at British Petroleum (chapter 4) is the means by which economic theory is reconsidered in the light of the concerns facing practitioners in order to produce the theory sought. In the process, it has been possible to draw on economic theories in an attempt to understand and explain the different contractual arrangements observed. The same theoretical framework, presented in a more lay language, served the interest of BP management in gaining a deeper and more general understanding of what is involved in setting up the governance of IT provision.

The multitude of IT services served by various contracts at BP, in conjunction with appropriate theoretical arguments, inspired the notion of service clustering. The characteristics and interrelationships among IT services provide opportunities and restrictions for clustering services under the same governance structure. The main benefit of having large clusters with many services is that problems of technical and managerial coordination are minimised. On the other hand, smaller clusters of homogenous services afford more specialised governance arrangements. In addition, although all contracts studied at BP were 'outsourcing' contracts, they differed in many significant ways. The relevant observations also corroborated the notion of the governance continuum.

The S-CAGE framework (chapter 5) brings together the clustering and the governance continuum and elaborates on their determination and inter-dependence. For example, when many diverse services are clustered together, the governance mechanisms are analytically similar to those found in internal IT departments. The composition of governance is presented in terms of three layers. This classification is derived from the study of contractual practices at BP and serves to facilitate understanding of the many issues involved and to direct attention to all aspects and implications of governance in practical circumstances.

While chapter 5 presents S-CAGE with reference to BP and the theory, chapter 6 attempts to re-address its relevance in alternative empirical contexts. It is shown that the framework is not just a convenient classification of practices at BP nor simply an exercise in rearranging economic theories. The contracting practices at ICI and Anglian Water have been made intelligible using this framework. More importantly, by using S-CAGE it is possible to explain differences in the governance structure between different companies or contracts. These two cases also showed that S-CAGE is most applicable at that management level where the organisational and contractual details are being set up.

Having recapitulated the emphasis in each previous chapter, the next section pinpoints the achievement of this thesis and its limitations, including a brief assessment of its validity. Some directions for further abstract and empirical research are discussed in section 3. Section 4 explores the question of whether S-CAGE is applicable to business activities other than IT. The chapter, indeed the thesis, concludes with a final brief illustration of the usefulness of S-CAGE.

2. Contribution and limitations

The main deliverable of this thesis is the S-CAGE framework. As discussed in chapters 3 and 5, it is submitted both as a theoretical basis for explaining the occurrence of various contractual and organisational configurations in the provision of IT systems and services, as well as a practical means for assessing and changing the governance of IT in actual commercial situations.

S-CAGE addresses many of the gaps identified in the IS literature. Specifically, the main criticism raised in chapter 1 is that some extant theories are too narrow either in the dimensions of governance and/or in the alternative outcomes. For example, attempts have been made to combine transaction and agency cost considerations or to introduce elements of risk or strategy. Outcomes are also restricted to either discrete forms of governance (e.g. market versus hierarchy or classical versus

relational contracting) or to variations in the variables (i.e. dimensions of governance) they study. These attempts are limited by the analytical or empirical methods they employ, or by the desire to produce a simple framework for decision making. The S-CAGE framework draws on a much broader theoretical canvas and attempts to synthesise the various insights available under a unifying framework. As a result the dimensions of governance are classified in three generic layers, each of which is analysed into much finer detail. In terms of outcomes, common categories of governance, such as systems development contracts or total outsourcing, are shown to be useful abstractions but nonetheless inadequate in characterising and analysing actual practices. An important theoretical contribution of S-CAGE is the argument that actual governance structures are theoretically infinite on a continuum between ideal pure markets and hierarchies and are determined by the exact configuration of the various elements of governance.

The S-CAGE framework is also consistent with the ontological and epistemological foundations laid down in chapter 3. The elements of governance, including, for example, performance measurement processes, explicit incentive schemes and decision making authority, are postulated as mechanisms that are expected to be in play but may not always be directly observable or some of them may dominate others in producing the observed effects. In any case, S-CAGE can be seen as a sourcebook of explanatory mechanisms from which the exact interactions and contingencies can be drawn in order to understand a given governance structure.

The incorporation of the realist ontology and epistemology in information systems research is another, secondary, contribution of the thesis.

Seen as such a sourcebook, S-CAGE can be used not only for explanatory but also for normative purposes. Specifically, S-CAGE can be used to assess and to make adjustments on an existing governance structure or to set up the governance on a

zero basis. Given a particular set of IT systems and services and a business context, the relevant elements of governance can be identified and configured accordingly.

The empirical work presented in chapters 4 and 6 has been instrumental in achieving the explanatory and normative goals of the proposed framework. Having said that, they are also a source of limitations. First, the organisations studied are large private companies in the UK[1] in the 1990s. Small and medium sized enterprises or public organisations are likely to face different constraints and opportunities and to be subject to different market forces, beyond the reach of the analysis in this thesis. Similarly, economic, legal and business systems significantly different to that found in the UK (e.g. in developing economies) may impinge differently on the capabilities and possibilities available to firms. Moreover, as social structures and business environments evolve over time, the validity and relevance of S-CAGE may be questionable in historical contexts or in the future.

Second, the empirical basis of the present work is relatively narrow. Although this does not necessarily compromise the validity of the substantive claims (see below), further refinements can be sought by using and assessing S-CAGE in more organisations. The two case studies in chapter 6 demonstrate that there is ample scope for refining and sharpening the framework by studying more contracts of different kinds.

Third, S-CAGE is not the ultimate, unifying theory of IT governance. It excludes important facets, most notably strategic considerations, supplier selection and negotiations. Although such issues do emerge in various parts of the analysis of governance elements and although S-CAGE can inform the relevant processes, they are not directly addressed.

---

[1] The fact that Anglian Water is a regulated utility was not found to have any repercussions beyond their retention of IT assets, as noted in chapter 6.

Fourth, S-CAGE is limited by the scope of the theories it encompasses and by the degree to which they are elaborated. The abstract theoretical process of extending the scope and refining the detail of the framework cannot be separated from further empirical investigation.

Finally, the whole thesis and the S-CAGE framework adopt the point of view of the buyer and user of IT services. The behaviour of suppliers and the dynamics of the supply market are assumed to be given. Among the antecedents of the choice of IT governance that have been omitted from the S-CAGE framework (but have been considered in chapter 1) is the political and commercial pressure that suppliers exert upon IT directors and CEOs in order to win a contract. With few exceptions (e.g. IDC and WBS, 1994) this is the perspective taken by the majority of the literature on the topic. The evolution of the supply side and the dynamics of its interaction with buyers (users) is worth significantly greater research attention. The importance placed by the Group IT Manager of ICI on the role of the IT industry (see chapter 6) is another hint that more work is needed in this direction. A preliminary idea in this respect is considered in section 3.1 further below.

2.1. A note on validity

Some further qualifications on the validity of the S-CAGE framework are worthwhile at this point. Having developed the framework, how can it be evaluated and on what criteria? The epistemological position adopted in this thesis is outlined in chapter 3. Briefly, the proposed theory is acceptable as long as it provides satisfactory explanations and guidance and until it is superseded by another theory (new or extant). Acid tests of predictive power and falsifiability have been dismissed on the basis of their principle.

Chapter 5 may have not satisfied those readers expecting to find a contingency framework. A pure contingency approach could provide a basis for (positivist) validation and an easier guide for practical decision making. As explained in

chapter 3, section 6.2, the temptation to create a contingency framework has been resisted. The configuration of governance is seen as complex and context-specific so that it cannot (and should not) be foreclosed by a small number of "if...then" clauses. Instead, it is argued, what is needed is an understanding of the operation and interaction of the various (structure and process) components of governance. This understanding is sufficient for both explanation and actual decision making.

At this point, it is worth reviewing how the perspective in this thesis responds to the widely held positivist criteria of validity (Yin 1989, Cavaye 1996). Internal validity is concerned with eliminating alternative and competing explanations for the same evidence. The realist epistemology accepts the possibility of sustaining complementary, even if conflicting, explanations for the same phenomenon. Examples include Newtonian physics, quantum mechanics and relativity theory in physics. All three theories coexist, although the last two are still incommensurable. In the social sciences there are several well-documented paradoxes, such as the paradox between social structure and human agency. Of course, all theories are subject to criticism and replacement by superior theories. In this thesis it is recognised that the S-CAGE framework excludes several important factors that determine the clustering of services and the choice of general organisational mode. By being inclusive in terms of the economic theories adopted, the possibility of competing explanations has been reduced. However the possibility that such theories may exist, does not invalidate the theoretical framework presented here, unless the alternative theories encompass the explanations provided within S-CAGE.

Construct or convergent validity ensures that the theoretical instrument reflects the concept being studied and not something else. No measurement instrument is being developed here. Nonetheless, it is worth emphasising once more that the thesis focuses on the elements of contracts and not on other facets of IT outsourcing relationships.

External validity refers to the generalisability of the framework. Clearly, the small number of empirical cases places a limit to the generality that can be claimed for S-CAGE. Having said that, there are three reasons supporting its general applicability. First, epistemologically, generalisation resides in real mechanisms and not in empirical regularities (chapter 3, section 6.1, paragraph 2). Second, the case study at BP offered evidence from at least three different contracts and a multitude of diverse opinions on outsourcing. Therefore, it is not a unitary datum of evidence. Third, the S-CAGE framework inherits the generality of the economic theories reviewed in chapter 2.

The criteria of reliability and replicability require that if the same project were to be repeated the conclusions would be the same. It is worth reiterating here that this criterion can be rejected in the social sciences due to the reflexive nature of human agents, who use their knowledge (including scientific theories) to reconstruct their reality (see chapter 3, section 3 and Numagami, 1998). In other words, theories can be self-fulfilling or self-negating. Having said that, the consistent application of the S-CAGE framework to the three companies is some evidence of reliability.

Finally, the acceptance of the framework by management at BP contributes not only to its practical value, but to its theoretical 'quality' as well.

3. Further research directions

Further development of the S-CAGE framework can be sought in three directions. First, in extending its empirical grounding, second, in extending its scope and third, in refining its analytical components.

In terms of empirical work, more intensive research in the form of case studies and action research is warranted. Additional observations from IT contracts can be a valuable source of further refinements. Further research should start by approaching major IT vendors and seeking access to their contracts.

A more profitable and exciting endeavour would be to use S-CAGE in changing the governance of IT in actual circumstances, especially in conjunction with alternative frameworks. Such work will clearly provide a better (comparative) assessment of the normative power of S-CAGE and will produce valuable lessons for practice. However, such a project would encounter serious problems of access to organisations.

The empirical work presented in previous chapters is limited to contracts of the commonly understood outsourcing kind. As a result, the S-CAGE framework has a certain degree of bias towards such contracts. Although these kinds of contracts constitute a rich array of interesting phenomena that are of contemporary interest to practitioners, further empirical work should place a more balanced emphasis on alternative types of governance, such as internal IT departments, and more discrete contracts. It is anticipated that internal departments will yield mechanisms equivalent to those currently prevailing in S-CAGE, whereas more discrete contracts will encompass small subsets of the governance mechanisms considered here.

The scope of S-CAGE can be extended in two ways. One possibility is to broaden the number of explanatory mechanisms by either incorporating other extant theories or by devising new. Another possibility is to broaden the range of related issues addressed by the framework, particularly strategic considerations, supplier selection and negotiations. Conceivably, the S-CAGE framework as it stands may not be able to sustain such an expansion without compromising its coherence or its explanatory and normative power. If this turns out to be the case, it will either have to be supplanted with a superior alternative or, at least, be linked to complementary theories of such issues in a consistent manner. In any case, the limited attention to such matters at present is a significant limitation that needs to be corrected. An appropriate combination of abstract and empirical research will have to be designed.

Another issue that needs to be addressed, is the nature of the role of internal IT organisation given the many governance possibilities for organising IT provision. This issue is not within the scope of S-CAGE but a consequence of it. A possible implication of adopting a framework such as S-CAGE within an organisation is that the residual IT staff are administrators of IT governance rather than providers of systems. There is some early work being done on the residual IS department (Willcocks and Fitzgerald 1994d, Venkatraman and Loh 1994, Markus and Benjamin 1996) but the issue of what is the core value-added of the internal IT organisation warrants further investigation. An interesting strand of research can grow in this direction.

Two opportunities for refining the analytical components of the framework, are identified here. The first relates to the method of clustering services in practice. A rather simple approach is presented in chapter 5, according to which it is assumed that it is *a priori* desirable to have all services clustered together in order to minimise the costs of technical and organisational coordination. Then individual services are separated if there are reasons to do so. This method may be sufficient in some cases but not necessarily always. It gives priority to the benefits of coordination while temporarily ignoring that governance structures for small service clusters are better matched to the characteristics of services and therefore more efficient and effective. Alternative, more sophisticated, clustering methods could take into account both criteria simultaneously, with the aim of minimising the number of service clusters while maximising each cluster's homogeneity. Such a method would probably require a detailed model of the criteria to be used as well as an elaborate optimisation technique. An even more advanced possibility would be a method of hierarchical clustering which would assign sets of services to hierarchically related suppliers (and respective governance structures).

These ideas point to the direction of a variety of possible approaches and techniques, including, for example, multi-criteria decision analysis. However, the

feasibility of developing formal quantitative models is questionable, on three accounts. First, the problem is quite complex, entailing many different technological and organisational parameters. Second, there are already significant difficulties in measuring empirically all the relevant parameters, such as asset specificity and uncertainty (see also chapter 2, section 8, paragraph 7). Third, the difficulties are compounded by the changing characteristics of the technologies, the user organisations and the IT vendors, over time. For these reasons, it is probably more worthwhile pursuing an incremental decomposition of the clustering decision, with a view to devising a practical problem-solving methodology rather than a formal algorithm. The purpose of this methodology would be to guide decision makers in clustering IT services and assigning them to internal or external IT providers. It would be a methodology that raises awareness of the issues among decision makers, rather than a technique that prescribes answers. The answers would be produced by informed decision makers, given the organisational context, rather than by an 'objective' algorithm.

The second opportunity for refining the analytical components of S-CAGE envisaged here entails the notion of governance feasibility, which is explored in some detail in the following section.

3.1. Governance feasibility frontiers

This section explores a series of conceptual ideas, based on the interaction between the governance continuum and the clustering continuum. These are abstract theoretical propositions, akin to much of the abstract theoretical development in economics, particularly the theories reviewed in chapter 2. However, appropriate research designs to develop and test these ideas encounter serious practical difficulties. Section 3.1.1 below, examines the feasibility of such research designs and methods.

We could postulate a surrogate measure for the governance continuum, ranging from 0, corresponding to the ideal market, to 1, corresponding to the ideal hierarchy. Williamson regularly uses asset specificity as such a surrogate (1996), although this is not necessarily the only determinant of governance, as discussed in chapter 2. The extent of contract discreteness and relationality (chapter 2, section 2.1) could be another surrogate for this purpose. If we were to survey the governance structure of given IT services among a large number of organisations, we could possibly derive the frequency distribution of governance structures for a given service. Williamson has also pondered on this issue (see chapter 2, section 8.1, paragraphs 7-8). Figure 7.1 shows three hypothetical examples of normal distributions.

The variance of the distribution would be a function of the heterogeneity of services and of other contingent factors determining the governance in each case. For example, the more standard and homogenous the service across organisations, the smaller the variance (i.e. narrower) of the frequency distribution (e.g. hardware procurement represented by distribution *i* in Figure 7.1). Other services such as software development (distribution *ii*) and desktop support (distribution *iii*) may differ significantly from case to case and, therefore, the respective governance structures would differ and the variance of the distribution would be larger.

The clustering continuum (Figure 5.2) can be redefined in terms of relative cluster size. If S is the total number of services and S_i the number of services in cluster i then relative cluster size is defined as S_i/S. This ratio takes the same values as the clustering ratio C/S (where C is the number of clusters), namely $(1/S, 1)$. However they have inverse meaning in that $1/S$ represents one cluster for all services in terms of the clustering ratio but only one service in cluster i in terms of relative cluster size. Accordingly, 1 represents one separate cluster for each service in terms of the clustering ratio but one cluster for all services in terms of relative cluster size.

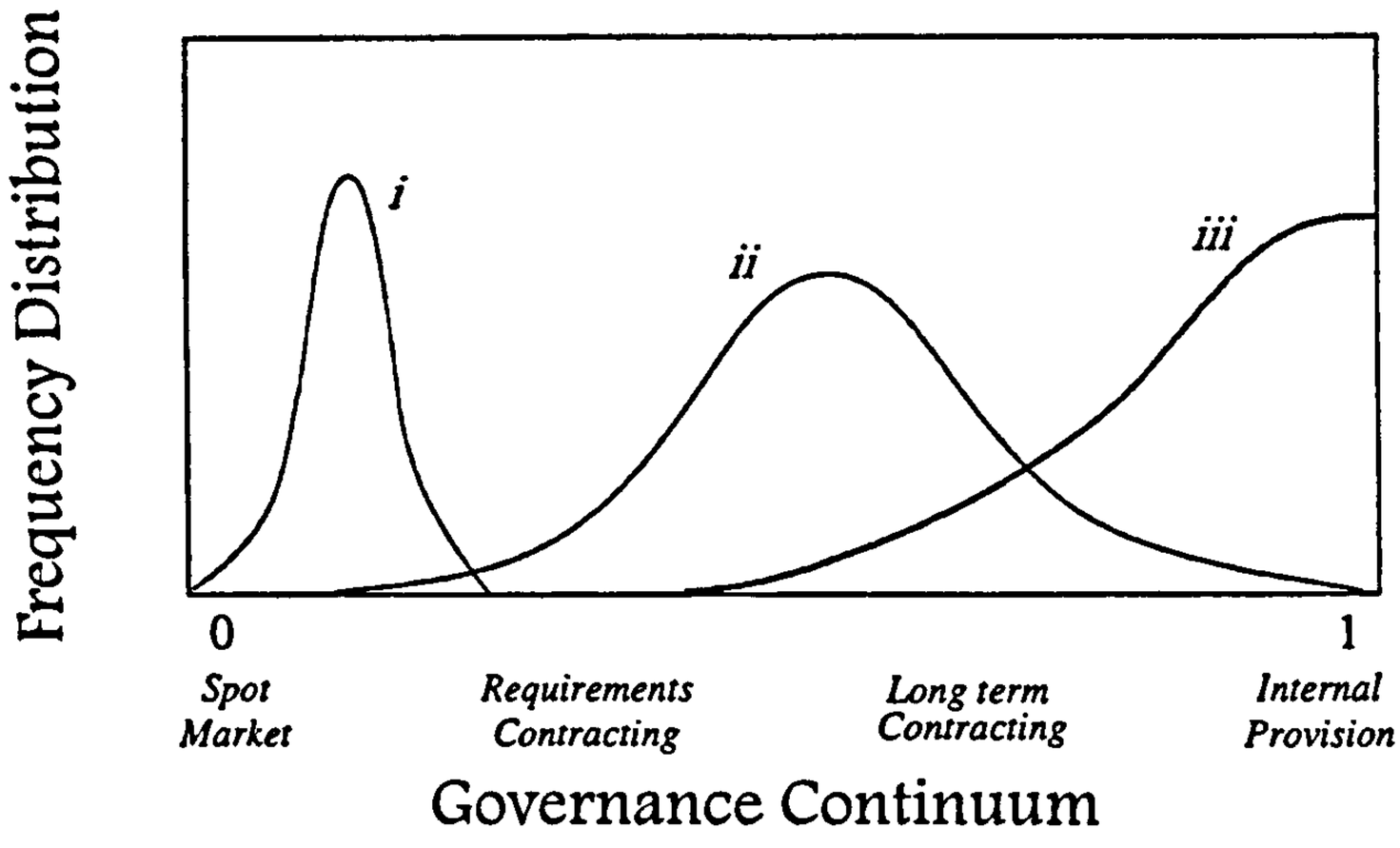


Figure 7.1

Frequency distributions of governance structures for individual services

These are hypothetical distributions. For example, they could correspond to hardware procurement (*i*), software development (*ii*) and desktop support (*iii*).

Assuming that an appropriate surrogate measure of the governance continuum can be devised, clusters could be plotted on a graph where the governance continuum is on the horizontal axis and the clustering continuum in terms of relative cluster size on the vertical. Figure 7.2 shows the example of total outsourcing towards the top of the clustering continuum (denoting a large cluster) and towards the right of the governance continuum (denoting a strongly relational long term contract) and the example of a software development contract denoting a small cluster under a more discrete contract.

On the same graph we can plot the whole governance portfolio allocation of an organisation at a given moment in time. Table 7.1 provides a hypothetical example of a portfolio of governance structures. This table lists six service clusters in terms of separate contracts. For each cluster, the allocation of services, budget, human resources (from the client's side), management support and customer/supplier involvement is then listed. The rightmost column shows graphically the extent of managerial involvement that each service cluster (contract) requires from the customer and the provider. The extent of this involvement is not necessarily commensurate to the amount of human resources allocated. Rather, it is a measure

of the extent to and proximity with which the two parties have to work together in managing service provision.

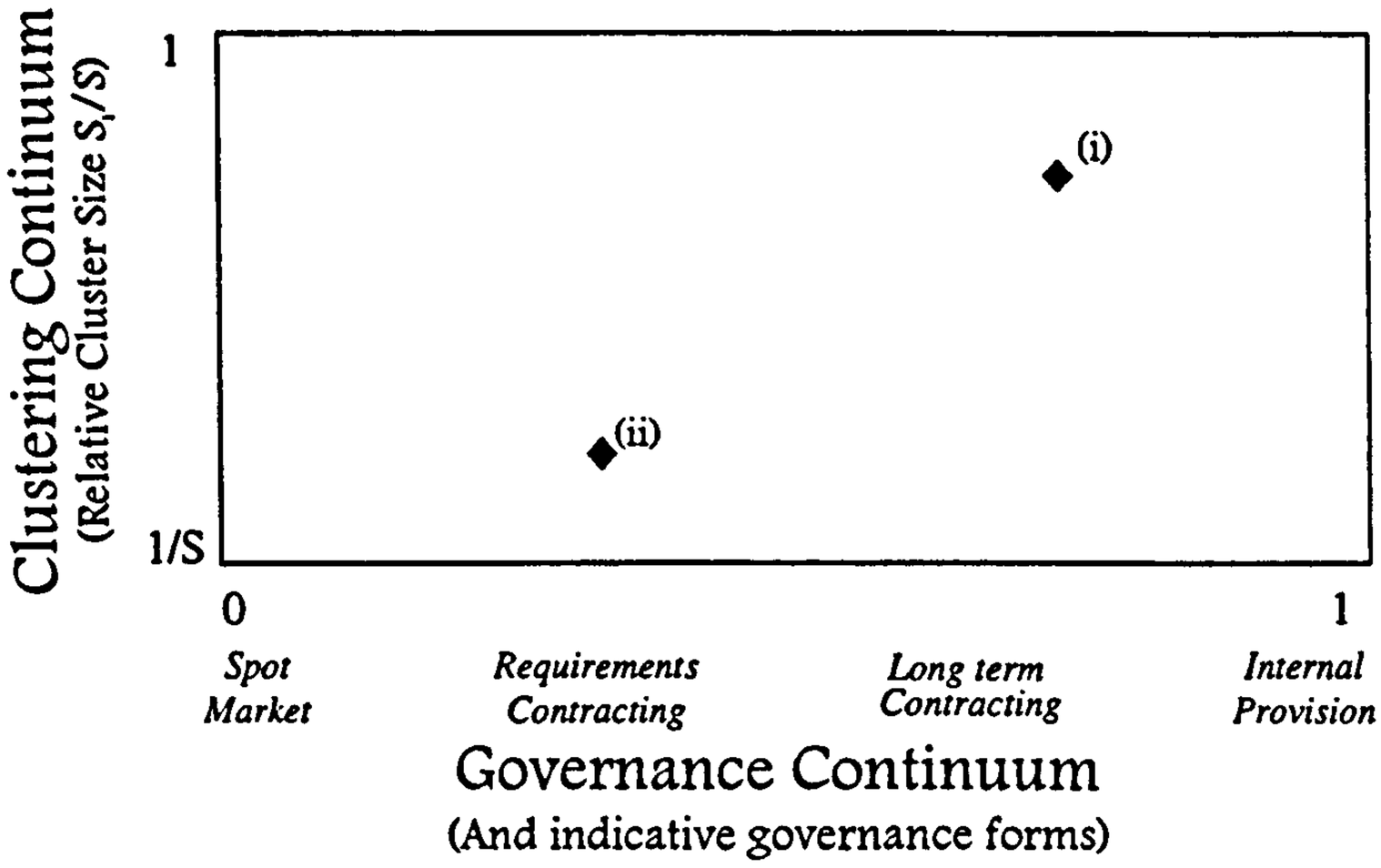


Figure 7.2

Examples of service clusters

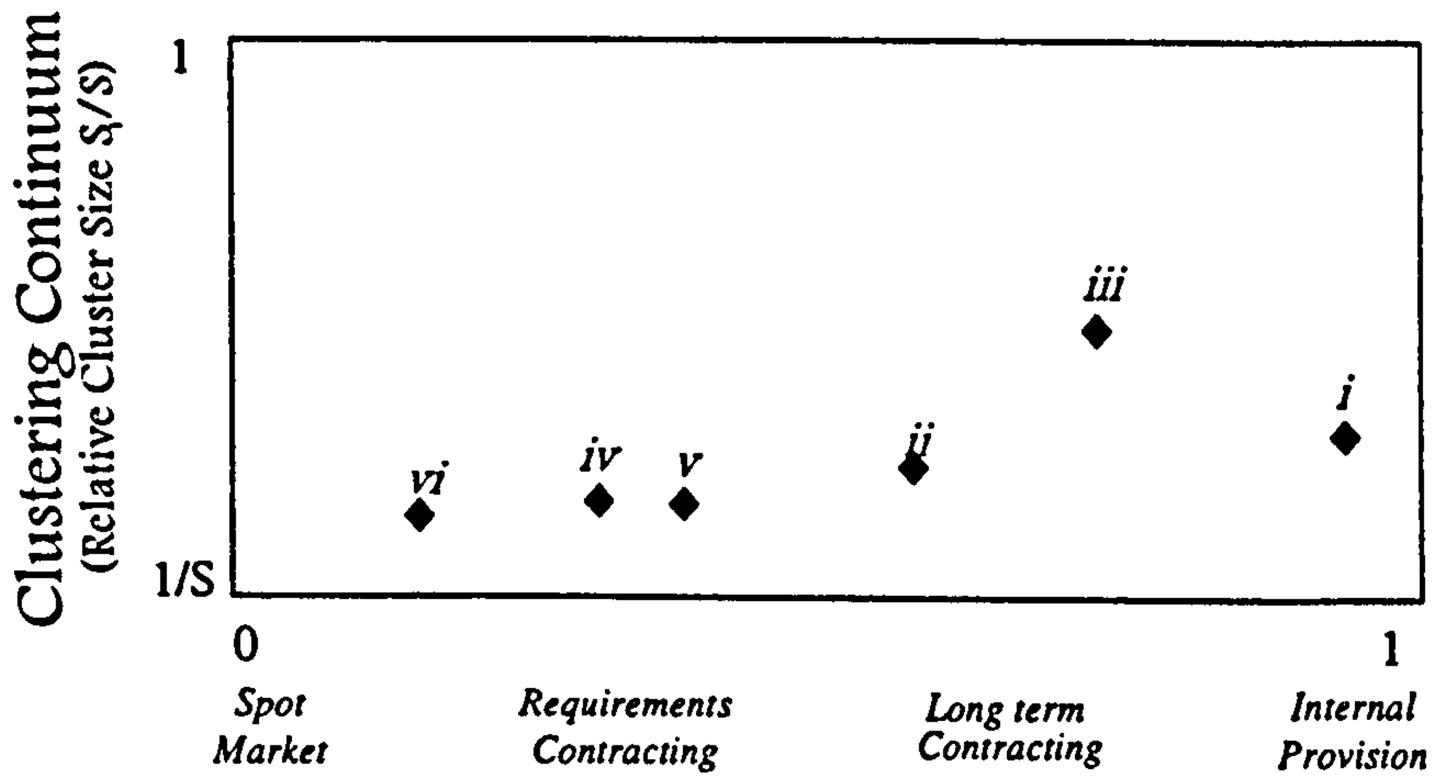
Two hypothetical examples. A total outsourcing contract (i) and a software development contract (ii). The position on the vertical axis denotes the size of the cluster. The positioning on the governance continuum denotes the extent of discreteness or relationality of the governance structure.

The hypothetical example of Table 7.1 is plotted in Figure 7.3. The internal IT department (rightmost point) corresponds to a fairly small cluster of services as it only includes strategy, planning, security and contract management. Outsourcing contract I includes fairly standardised services and can, therefore, afford a more market-like governance. Outsourcing contract II includes more company-specific services and thus requires more sophisticated governance with hierarchical elements. Project contracts A and B are typical systems development contracts. The hardware lease contract is the most discrete of all (leftmost point). This graph is another useful means of depicting governance portfolios[2].

[2] The graph is not drawn to exact scale. The total height of all points in the same governance portfolio must equal 1 ($\sum_{i=1}^{c} S_i/S = 1$).

	Governance Portfolio	Allocation				
		Services	Budget	Human Resources	Management Support	Management Involvement *Customer* ⟶ ⟵ *Provider*
i	Internal Department	Strategy, Planning, Security	5%	1 Senior Mgr 1 Mid. Mgr 1 Technician	Steering committee	⟶
ii	Outsourcing contract I	WAN Telecoms	15%	1 Contract manager	Service level monitoring Contract management	⟶|⟵
iii	Outsourcing contract II	Data Centres, LAN, Help desk	45%	1 contract manager	Service level monitoring Contract management	⟶|⟵
iv	Project contract A	System A	15%	1 Analyst 1/3 Contract manager	Monitoring and testing	⟶|⟵
v	Project contract B	System B	10%	1 Analyst 1/3 Contract manager	Monitoring and testing	⟶|⟵
vi	Lease contract	Hardware	10%	1/3 Contract manager	Contract management	⟶|⟵

Table 7.1
Governance Portfolio Allocation (hypothetical example)



Governance Continuum
(And indicative governance forms)

Figure 7.3
Governance Portfolio Allocation
This graph maps the hypothetical clustering of Table 7.1 .

It is also possible to demonstrate on this graph the theoretical argument that larger (and hence more heterogeneous) service clusters require more sophisticated governance structures, of the kinds located towards the right of the governance continuum. Relatively more discrete (market-like) transactions can handle the procurement of services that are relatively more standard with little or no asset specificity, uncertainty and measurement problems. As clusters grow in size, such sources of transaction costs are likely to compound thus making discrete transactions infeasible. All this means that clusters on the left of the governance

continuum are necessarily more limited in size as compared to clusters organised under governance structures towards the right of the governance continuum.

At the extreme, it can be shown that it is impossible to procure all IT services that an organisation makes use of, under a single discrete contract (such as off-the-shelf purchasing). It is similarly impossible for any company to produce all services, including all hardware and software, in-house. Thus, clusters on the right cannot include all services (i.e. their maximum relative cluster size is necessarily less than 1). Figure 7.4 postulates a convex governance feasibility frontier which represents this limit to the maximum feasible size of clusters over the range of the governance continuum.

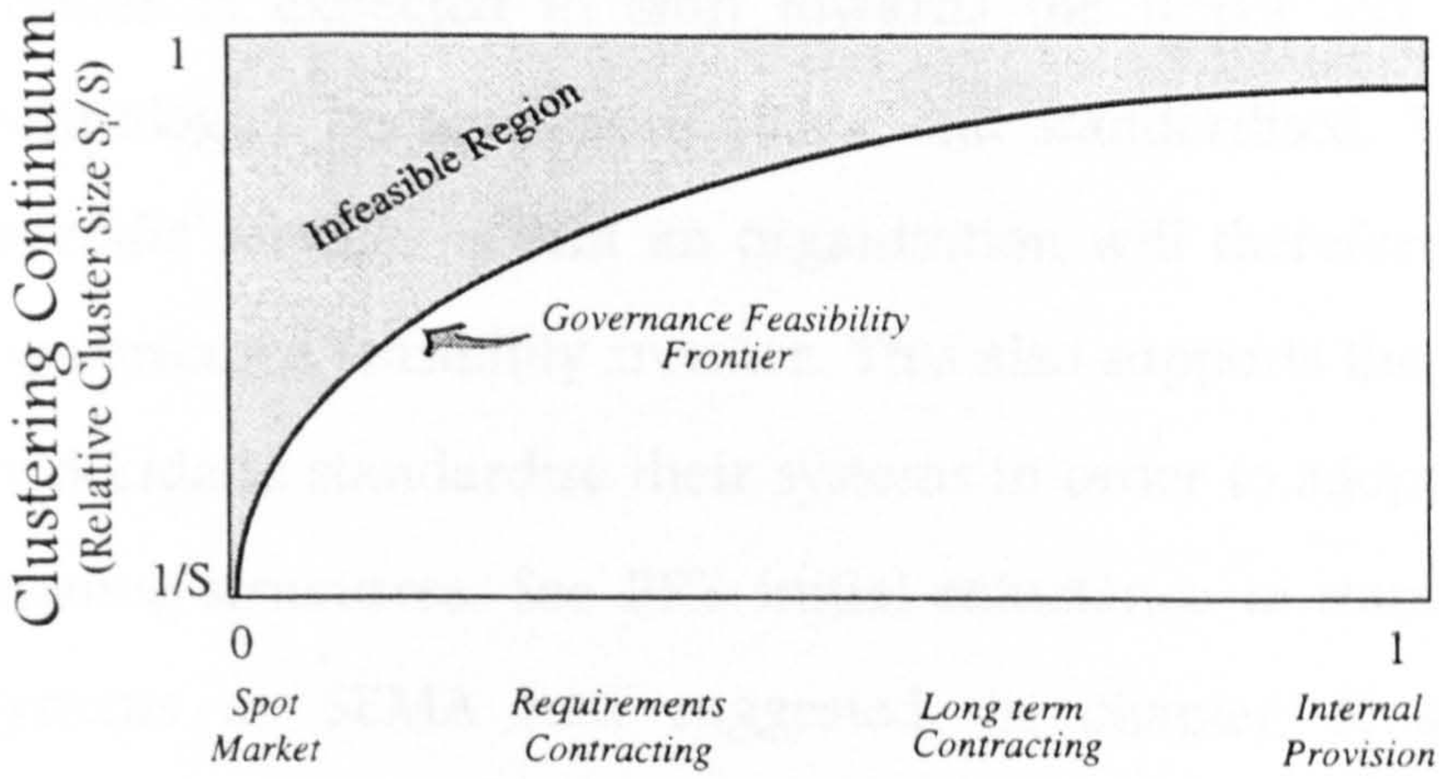


Figure 7.4
Governance Feasibility Frontier

The governance feasibility frontier represents the extent to which more or fewer IT services (larger or smaller clusters) can be organised under a single governance for each type of governance. It can also be thought of as an efficient frontier. Given a cluster of services, if their governance requirements can be measured using the hypothetical surrogate mentioned above, they should be organised under the leftmost feasible governance structure[3]. The efficient governance thus obtained mitigates market transaction costs while eschewing unnecessary bureaucratic slack.

[3] The implicit assumption here is that the respective markets exist and that the law sustains the desired governance structures.

The shape and the location of the governance frontier are currently unknown and possibly indeterminate. The shape in Figure 7.4 is indicative. The governance feasibility frontier may vary with time, industry and firm and its shape and location are expected to be affected by the nature of IT systems and services and by market conditions.

New technologies and innovative systems are likely to be associated with significant uncertainty, specificity and performance measurement problems. Therefore, under such conditions the frontier is expected to shift towards the bottom right of the diagram. In contrast, as technologies and services mature over time, such as data centres and help desks, they are less subject to such conditions. As a result, the governance frontier is expected to shift towards the upper left corner of the diagram, as technologies become more stable and standardised. The mixture of standard and specific services within an organisation will therefore determine, to an extent, the governance feasibility frontier. This also supports the argument that companies may decide to standardise their systems in order to adopt more market-oriented governance structures. See BP's initial reluctance to standardise certain information systems as SEMA had suggested, in chapter 4, section 8, last paragraph.

The growth of the market for different systems and services is also another determinant of governance feasibility. The governance possibilities faced by companies in a remote region or in developing economies are expected to be fewer (and perhaps even qualitatively different) to those available in large developed markets. Market growth is expected to shift the governance frontier towards the upper left corner of the diagram.

This reference to the IT supply market points to the fact that the preceding discussion is made from the point of view of an individual user organisation (i.e. the demand side). Indeed, this is the perspective taken by the whole thesis. However, a symmetric argument can be articulated for the supply side. The IT services market is expected to offer smaller and homogenous clusters of services

under relatively more discrete contracts and larger, heterogeneous service clusters under more relational governance structures. Clearly, internal organisation (hierarchy) amounts to vertical integration, which for the purposes of this argument must be seen as a discrete (as opposed to continuous - see end of chapter 2) change in governance structure.

A possible supply-side governance frontier is depicted in Figure 7.5 alongside a demand-side governance frontier. The former demarcates the minimum cluster size sustainable by a given governance structure and/or the maximum governance requirements for an IT service cluster of given size. Accordingly, the latter demarcates the maximum cluster size for a given governance structure and/or the minimum governance requirements for a cluster of given size. The area between the two curves can be seen as the range of feasible governance structures for IT service clusters that should be expected to be found in practice.

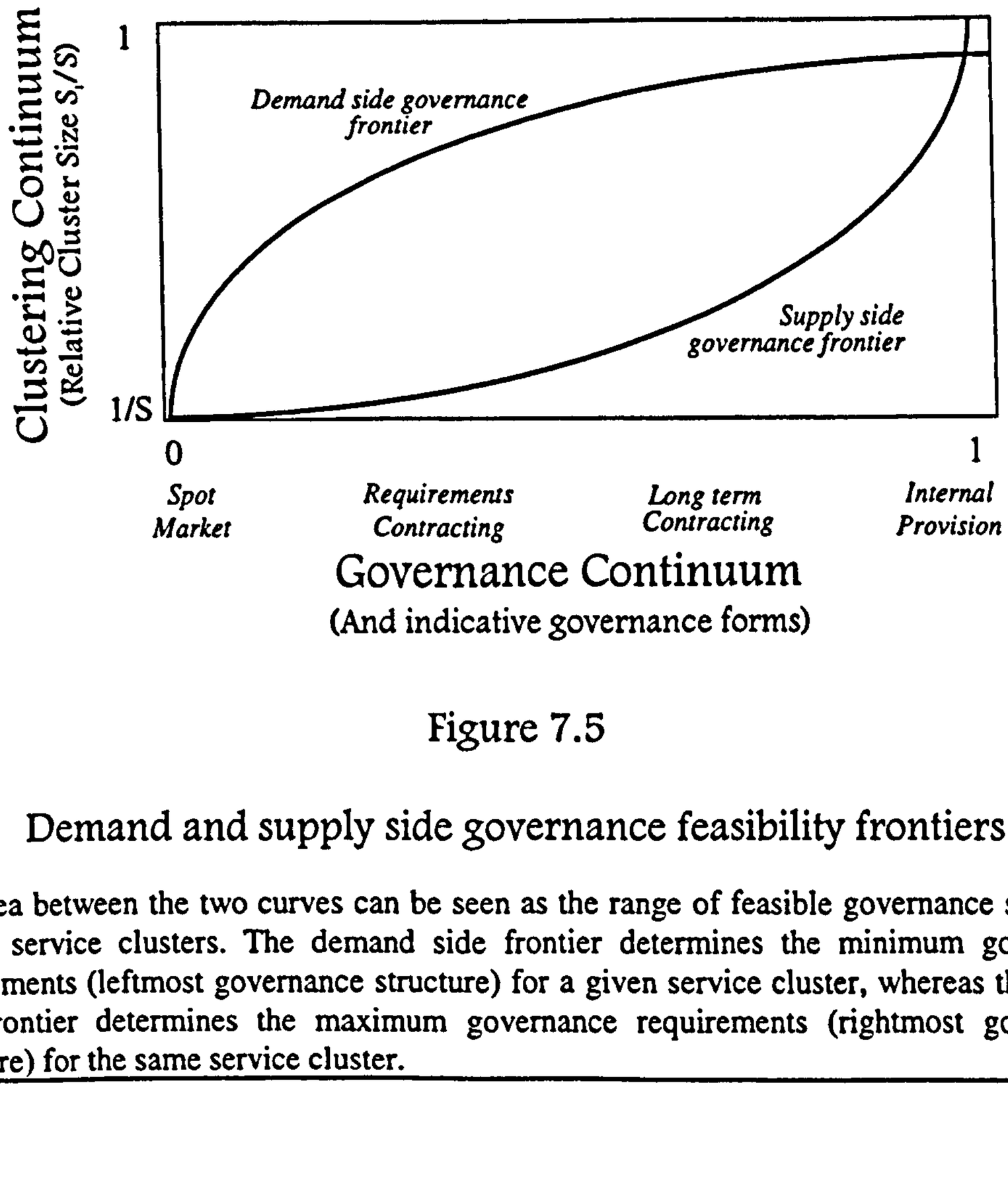


Figure 7.5

Demand and supply side governance feasibility frontiers

The area between the two curves can be seen as the range of feasible governance structures for IT service clusters. The demand side frontier determines the minimum governance requirements (leftmost governance structure) for a given service cluster, whereas the supply side frontier determines the maximum governance requirements (rightmost governance structure) for the same service cluster.

3.1.1. Research designs for governance feasibility

Earlier suggestions to extend the empirical grounding and theoretical scope of the S-CAGE framework have an immediate practical benefit in terms of strengthening the theory and enhancing its usefulness. In contrast, the notion of governance feasibility frontiers is an abstract, conceptual extension of the notions of the clustering continuum and the governance continuum. As it stands, it is more of a definitional proposition than a testable hypothesis. This approach is common in both traditional as well as recent institutional[4] economic theories[5] (Jensen, 1983). For example, in a similar vein, Riordan and Williamson (1985) and Williamson (1996:108-109) draw governance cost curves.

As Jensen (1983) argues extensively, the potential value of such abstract, definitional, propositions is that they can stimulate debate, enhance our understanding of the issues and, eventually, lead to the development of partial, incremental, testable hypotheses. In this sense, the empirical estimation of the governance feasibility frontier for any industry, organisation or technology is not the point. Besides, it would be very difficult, if not impossible to do in practice. On the other hand, understanding what makes one or more IT services more or less amenable to contracting under different governance structures, is both more attainable (if carried out incrementally) and more valuable for practitioners.

Empirically deriving the determinants of governance feasibility for a given set of IT services within a given organisational context, could be a fruitful research direction. The literature on IT outsourcing and the empirical work carried out for this thesis, show that one issue practitioners are deliberating is whether IT is, or will ever be, suitable for 'simple' purchasing. For example, a senior BP manager raised this possibility (see chapter 4, section 4, paragraph 6). Moreover, common

---

[4] Williamson (1985) coins the term 'new institutional economics' to refer collectively to transaction cost and related theories. See chapter 2, Figure 2.1 and related text.

[5] See also Figure 3.1 and related text in chapter 3.

understanding suggests that industry maturity with a technology can be an enabler of more hands-off contracting practices. Clearly, understanding the determinants of governance feasibility has both theoretical and practical value.

Both intensive and contextual, as well as extensive and statistical research approaches are suitable for this purpose. As mentioned in chapter 2, section 8, the study of governance structure entails significant measurement problems. The theoretical constructs with which we are dealing in these theories are not directly observable in practice in the way that quantities and prices are. However, it is still worthwhile to try to develop appropriate measurement instruments based on indirect manifestations of the mechanisms of interest. Such endeavours have been met with some success in other areas of information systems, such as service quality (e.g. Pitt et al 1995) and information privacy (Smith et al 1996). Empirical research of the transaction and agency determinants of organisational form, abound in economics, marketing and organisational studies. Quantitative research on this subject aims at developing suitable measurement instruments that approximate the constructs of interest (such as asset specificity and risk aversion) satisfactorily.

In contrast, qualitative research relies on interpretive evidence of the presence and magnitude of such constructs. As suggested in chapters 4, 5 and 6, the presence and magnitude of asset specificity or uncertainty, is not always objectively understood. It often depends on the perceived value of IT assets and on subjective expectations about the future of the technology and of the organisation. Therefore, further research on the determinants of governance feasibility ought to combine both quantitative and qualitative sources of evidence, in both intensive and extensive research designs (see footnote 5 above).

A final epistemological note is warranted at this stage. On the whole, the proposals advanced in this subsection appear to linger towards the empiricist research tradition that has been rejected earlier in chapter 3. Transcendental realism does

not make any a priori judgements for or against particular techniques for empirical research. It does, however, restrict the types of theoretical claims that they sustain. For example, there is nothing inherently wrong in developing an empirical instrument for measuring the idiosyncrasy of governance structures, as long as we qualify the result as representative of experiences and not of real structures and mechanisms. Such measures of experience are still important elements of realist research. The epistemological assessment of the above proposals should become an integral part of their substantive development.

4. Information technology and other business activities

A question that naturally arises is to what extent the S-CAGE framework is specific to the procurement of IT or of general applicability to any business activity. The straightforward answer is that it has been developed in the context of and for the purposes of IT procurement and therefore, any claims of broader generality cannot be sustained by this research. Surely though, there are good reasons to expect that this framework should be valid outside IT.

The question asked can be restated as to what extent IT is similar or different to other business activities. The argument made here is that the provision of IT services presents certain distinct differences. In particular, IT permeates all parts of the organisation, or, in other words, it is not organisationally contained. Not only is IT found in every part of the business but it is also intrinsically linked with it. While computers, software and communications are typically standard products, the services they perform are typically inseparable from the business activity they serve. Every change in organisation and management has to be followed by appropriate adaptations in the respective information systems and *vice versa*[6].

---

[6] One implication of this argument is that inasmuch as information systems are inseparable from the business itself, the expectation that IT suppliers should recommend innovative uses of IT, amounts to expecting innovations to the core business from an IT supplier. This is a particularly problematic issue, as discussed in chapter 5, section 4.2, paragraphs 5-8 and in the next section in this chapter.

This is not the case with other business activities. For example, the procurement of manufacturing designs and components is a complex process, often involving a large number of services and (sub)systems and significant asset specificity. However, it is organisationally isolated and requirements can be specified with engineering precision. Warehousing and distribution is a function that may involve significant uncertainty but, again, it is organisationally separated and well defined. Similar arguments can be made for most business activities that are candidates for contracting out, from accounting/finance and advertising to catering and mail facilities.

Further implications are manifested when trying to specify contractual requirements. Requirements are ill-defined and subject to frequent change. The same is true for support services. User organisations often wish to specify requirements in business terms in order to distance themselves from technological issues. Sometimes, though, contracts of useful duration are not sustainable if business requirements change regularly and are not measurable. Faced with this problem, companies have to resort to specifying requirements for technical resources (inputs). But then, in the face of technological innovation and changing market prices, they risk being locked into obsolete or too expensive systems and/or skills.

To address these difficulties, the S-CAGE framework places particular attention in the problems related to service definition and requirements specification under the first layer of governance (assurance of service delivery). In addition, innovation is facilitated with appropriate incentives (elaborated under the second layer of governance) and, more importantly, through coordination processes that communicate relevant information to the appropriate persons in the organisation (discussed under the third layer of governance).

To be sure, not all IT systems and services are subject to such problems to the same degree. This, in turn, points to another distinct feature of IT, namely, the multiplicity, heterogeneity and interconnectedness of systems and services. The notion of clustering aims at providing a basis for balancing the benefits of specialisation that accrue to separated organisation of services and the requirements for coordination.

The only other business activity that has to coordinate a diverse range of interrelated components is production, for instance in the clothing and car industries. In these industries, however, the division of clusters is given by the designs, the technology and the structure of the market. IT services are not separable with the same ease and their interfaces are less well defined. Meanwhile, the market is not segmented into specialist suppliers, with few exceptions such as telecommunications.

To summarise, the S-CAGE framework is geared towards IT procurement. As such, it emphasises certain distinct features of IT. However, all its theoretical claims must have broader applicability since they are based on general economic theories. Therefore, S-CAGE must be relevant to other business activities, even if it may not give a balanced emphasis on the issues that are important in every such case.

5. A final illustration

IT outsourcing is still a topical issue today, even though it ranked last (20th) in the 1994 survey of IS management issues in the US (Brancheau et al 1996). Figure 7.6 shows the number of articles (academic as well as professional) related to IT outsourcing that have been recorded in the ABI/INFORM bibliographic database over the period 1988-'96. Even if these figures represent a lot of industry hyperbole, they are an indication of an increasing interest in the topic.

Some of this interest may be a result of the continuous growth of the industry (see the introduction of chapter 1) or from concerns that arise as experience accumulates from earlier contracts. Such is a recent article[7] reporting on the initiatives of the International Information Technology Users Group, an exclusive group of the 20 largest IT outsourcing users in the world. The aim of this group is to find solutions to a range of problems that arise in outsourcing relationships, mainly due to contractual incompleteness. Of course, this is not the term used in the article. Instead, the article refers to 'soft promises' which include, among others, "prompt responses to letters and telephone calls, critical technological failures beyond the contracted targets, unexpected changes in staff or failure to meet unexpected changes in demand". There is also reference to minor opportunistic tactics by both customers and suppliers and to trust which is defined as 'no unpleasant surprises'. Finally, the largest outsourcing users stress their expectation that suppliers deliver "*value added service which should be a new, specific and innovative IT-led creation that increases profits*".

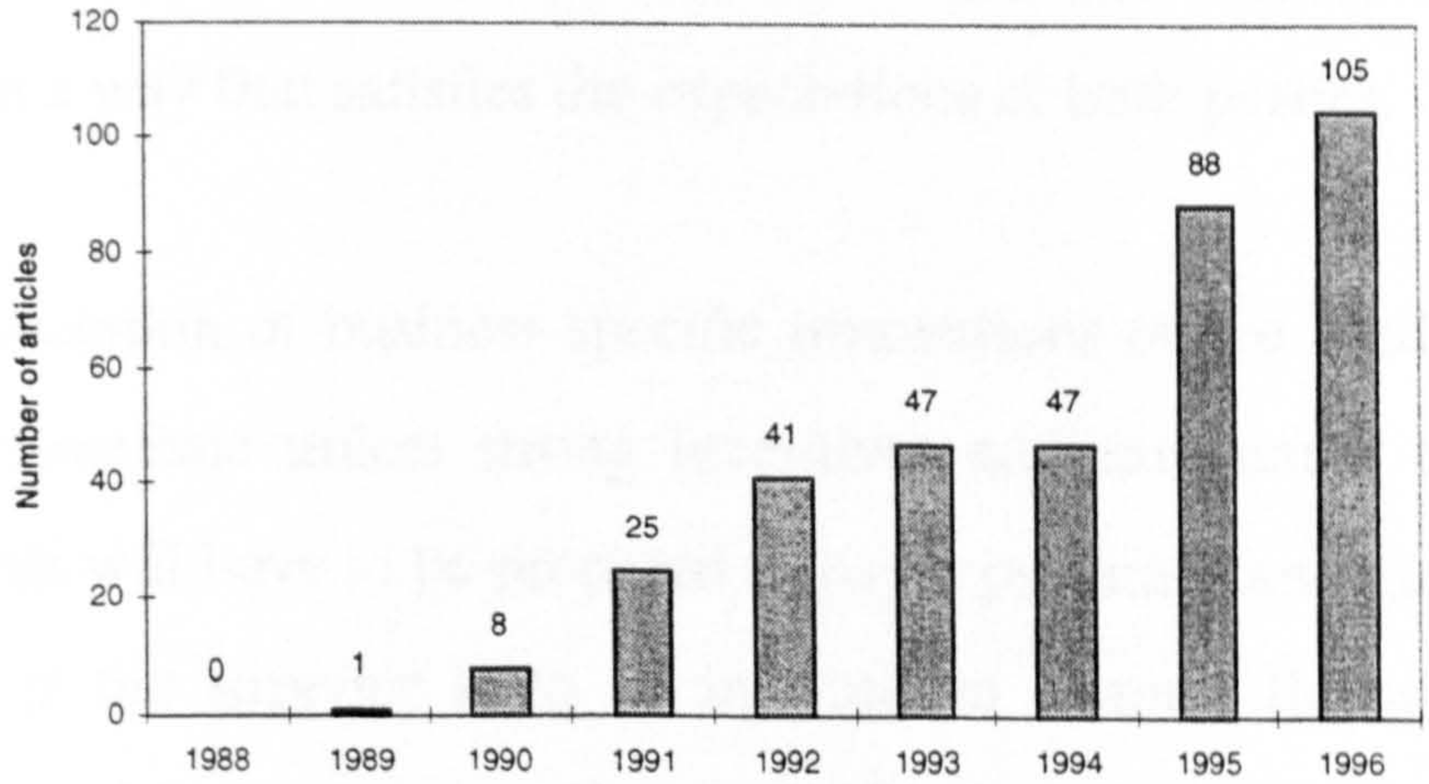


Figure 7.6

Number of articles on 'information technology outsourcing' in the ABI/INFORM bibliographic database, over the period 1988-1996

If this report summarises the main problems that arise in big outsourcing deals, the S-CAGE framework can help separate the issues in order to understand them from a clearer perspective. Breaches of the so called soft promises can be limited by

[7] By Tony Collins, *Computer Weekly*, Thursday 31 July 1997, pp. 1, 16-17.

introducing more detailed requirements and measures of performance in the contract. Indeed, the above mentioned user group tries to devise scorecards for use by both client and supplier staff with a view to recording such breaches from both sides. However, such breaches are ultimately unavoidable on account of the necessary incompleteness of contracts. Moreover, the difficulty (cost) of devising and administering additional metrics may not be worth the marginal improvement in the relationship (see the SLR lifecycle in chapter 5, section 4.1.2.1), although, admittedly, it may be worth doing so collectively within a consortium of clients and suppliers such as the above mentioned group.

The treatment of trust in S-CAGE suggests that unless there are severe sanctions against minor opportunistic tactics, their short term benefits will always exceed the respective 'costs' in terms of unpleasant friction (since they never threaten the continuation of the contract). Moreover, trust has to be calculative and rigorously supported by appropriate governance mechanisms. The terms trust and partnership should not be synonyms for lousy contracting. The notion of trust as 'no unpleasant surprises' points directly to the importance of a common culture for addressing contingencies in a way that satisfies the expectations of both parties.

Finally, the expectation of business-specific innovations of the kind mentioned in the article is unrealistic unless strong incentives and safeguards are offered. In particular, clients will have to be prepared to pay a premium and guarantee a long term contract, if the supplier is to be induced to commit the unique, specific investment required for such innovations. Suppliers are furthermore deterred from such activity because the client cannot commit not to hire key supplier personnel when the latter acquire business-specific expertise. As it turns out, anecdotal evidence suggests that vendors purposefully rotate key staff in order to pre-empt this possibility. If client organisations are not prepared to (or cannot) offer the necessary incentives and safeguards, they will have to develop and maintain appropriate skills and organisational structures to produce such innovations internally. As mentioned in the previous section, this specificity of innovations in IT

to the core of the business is a characteristic that distinguishes IT from other business activities.

The purpose of the S-CAGE framework is to provide a theoretically grounded basis for illuminating practical problems, such as this brief exposition, with a view to supporting further decision making. There remains ample scope for broadening and strengthening our understanding of the effective and efficient governance of IT service provision. It is hoped that further research activity will take on this challenge.

REFERENCES

Ahituv N, S Neumann and M Zviran (1989), "Factors affecting the policy for distributing computer resources", *MIS Quarterly*, Vol. 13, No. 4, pp. 389-402.

Alchian AA and H Demsetz (1972), "Production, information costs and economic organisation", *American Economic Review*, Vol. 62, pp. 777-795.

Allen BR and AC Boynton (1991), "Information architecture: in search for efficient flexibility", *MIS Quarterly*, Vol. 15, No. 4, pp. 435-445.

Allen DW and D Lueck (1995), "Risk preferences and the economics of contracts", *American Economic Review*, Vol. 85, No. 2, pp. 447-451.

Alpar P and A Saharia (1995), "Outsourcing information system functions: an organization economics perspective", *Journal of Organizational Computing*, Vol. 5, No. 3, pp. 197-217.

Alston LJ and W Gillespie (1989), "Resource coordination and transaction costs: A framework for analysing the firm/market boundary", *Journal of Economic Behavior and Organization*, Vol. 11, pp. 191-212.

Altinkemer K, A Chaturvedi and R Gulati (1994), "Information systems outsourcing: issues and evidence", *International Journal of Information Management*, Vol. 14, pp. 252-268.

Anderson E and RL Oliver (1987), "Perspectives on behaviour-based versus outcome-based salesforce control systems", *Journal of Marketing*, Vol. 51, pp. 76-88.

Ang S (1993), "The etiology of information systems outsourcing", *PhD Dissertation*, Department of Information and Decision Sciences, University of Minnesota.

Ang S (1994), "Toward conceptual clarity of outsourcing", in *Proceedings, IFIP TC8 Information Systems International Working Conference*, Queensland, Australia, May, pp. 131-141.

Ang S and C Beath (1993), "Hierarchical elements in software contracts", *Journal of Organizational Computing*, Vol. 3, No. 3, pp. 329-361.

Angell IO (1996), "The signs are clear: the future is inequality", *The Independent Newspaper*, Wednesday 25 September 1996.

Aoki M (1990), "The participatory generation of information rents and the theory of the firm", in Aoki et al (1990), pp. 26-52.

Aoki M, B Gustafsson and O Williamson (Eds.) (1990), *The firm as a nexus of treaties*, Sage.

Apte U and R Mason (1993), "Global disaggregation of information-intensive services", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Argyris C, R Putnam and D McLain Smith (1985) *Action Science: Concepts, methods and skills for research and intervention*, Jossey-Bass Publishers.

Aubert B, S Rivard and M Patry (1994), "Development of measures to assess dimensions of IS operation transactions", *Proceedings of the 15th International Conference on Information Systems*, 14-17 December, Vancouver, Canada, pp. 13-26.

Aubert B, S Rivard and M Patry (1996), "A transaction cost approach to outsourcing behavior: some empirical evidence", *Information & Management*, No. 30, pp. 51-64.

Aylott B (1995), "Three main elements of managing an outsourcing contract", *Best practice guidelines for outsourcing*, KPMG IMPACT Programme, London: HMSO.

Bakos JY and CF Kemerer (1992), "Recent applications of economic theory in IT research", *Decision Support Systems*, Vol. 8, pp. 365-386.

Banker RD and CF Kemerer (1989), "Scale economies in new software development", *IEEE Transactions on Software Engineering*, Vol. 15, No. 10, pp. 1199-1205.

Banker RD and SA Slaughter (1994), "Project size and software maintenance productivity: empirical evidence on economies of scale in sofware maintenance", Proceedings of the 15th International Conference on Information Systems, 14-17 December, Vancouver, Canada, pp. 279-289.

Barron T (1992), "Some new results in testing for economies of scale in computing: 1985 and 1988 data", *Decision Support Systems*, Vol. 8, pp. 405-429.

Barzel Y (1982), "Measurement cost and the organisation of markets", *Journal of Law and Economics*, Vol. 25, pp. 27-48.

Baskerville RL and AT Wood-Harper (1996), "A critical perspective on action research as a method for information systems research", *Journal of Information Technology*, Vol. 11, pp. 235-246.

Benbasat I, DK Goldstein and M Mead (1987), "The case research strategy in studies of information systems", *MISQuarterly*, September, pp. 369-386.

Benko C (1992), "If information system oouts is the solution, what is the problem?", *Journal of Systems Management*, Nvember, pp. 32-35.

Bhaskar R (1978) *A realist theory of science*, 2nd Edition, Harvester Press.

Bhaskar R (1979) *The possibility of naturalism*, Harvester Press.

Black G (1996), "Logica wins £27m deal to steer Ford's IT forward", *Computer Weekly*, 28 November, p. 16.

Blaug M (1980) *The methodology of economics or how economists explain*, Cambridge University Press.

Blois KJ (1972), "Vertical quasi-integration", *Journal of Industrial Economics*, Vol. 20, No. 3, pp. 253-272.

Blois KJ (1972), "Vertical quasi-integration", *Journal of Industrial Economics*, Vol. 20, pp. 253-272.

Brancheau JC, BD Janz and JC Wetherbe (1996), "Key issues in information systems management: 1994-95 SIM Delphi results", *MIS Quarterly*, Vol. 20, No. 2, pp. 225-242.

Brown CV (1997), "Examining the emergence of hybrid IS governance solutions: Evidence from a single case study", *Information Systems Research*, Vol. 8, No. 1, pp. 69-95.

Brown CV and SL Magill (1994), "Alignment of the IS functions with the enterprise: Towards a model of antecedents", *MIS Quarterly*, Vol. 18, No. 4, pp. 371-403.

Brown CV and SL Magill (1998), "Reconceptualising the context-design issue for the information systems function", *Organization Science*, Vol. 9, No. 2, pp. 176-194.

Buck-Lew M (1992), "To outsource or not?", *International Journal of Information Management*, Vol. 12, pp. 3-20.

Buttrick J (1952), "The inside contract system", *The Journal of Economic History*, Vol. 12, No.3, pp. 205-221.

Cash JI Jr. and PR Lawrence (Eds.) (1989), *The information systems research challenge Vol. 1: Qualitative research methods*, Harvard Business School Press, Boston.

Castells M (1996), *The information age: economy, society and culture*, Volumes I, II and III, Blackwell.

Cavaye ALM (1996), "Case study research: a multi-faceted research approach for IS", *Information Systems Research*, Vol. 6, pp. 227-242.

Chalmers AF (1982) *What is this thing called science?*, Second Edition, Open University Press.

Checkland P (1991) "From framework through experience to learning: the essential nature of action research," in H-E. Nissen, H.K. Klein, R.A. Hirschheim (Eds.), *Information Systems Research: Contemporary Approaches and Emergent Traditions*, North-Holland, pp. 397-403.

Cheon M, V Grover and J Teng (1995), "Theoretical perspectives on the outsourcing of information systems", *Journal of Information Technology*, Vol. 10, No. 4, pp. 209-219.

Clark CE, NC Cavanaugh, CV Brown and V Sambamurthy (1997), "Building change-readiness capabilities in the IS organisation: Insights from the Bell Atlantic Experience", *MIS Quarterly*, Vol. 21, No. 4, pp. 425-456.

Clark PA (1972), *Action research and organizational change*, Harper and Row: London.

Clark RC (1985), "Agency costs versus fiduciary duties", in Pratt and Zeckhauser (1985), pp. 55-80.

Clark T and R Zmud (1993), "The outsourcing of information services: tranforming the nature of business in the information industry", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Clark T, R Zmud and G McCray (1995), "The outsourcing of information services: transforming the nature of business in the information industry", *Journal of Information Technology,* Vol. 10, No. 4, pp. 221-237.

Coase RH (1937), "The nature of the firm", *Economica*, Vol. 4, pp. 386-405.

Coase RH (1993), "The nature of the firm: Influence", in Williamson and Winter (1993), pp. 61-74.

Cowen T and A Glazer (1996), "More monitoring can induce less effort", *Journal of Economic Behavior and Organization*, Vol. 30, pp. 113-123.

Cremer J (1990), "Common knowledge and the coordination of economic activities", in Aoki et al (1990), pp. 53-76.

Crocker KJ and SE Masten (1991), "Pretia ex machina? Prices and process in ong-term contracts", *Journal of Law and Economics*, Vol. 34, pp. 69-99.

Cronk J and J Sharp (1995), "A framework for deciding what to outsource in information technology", *Journal of Information Technology,* Vol. 10, No. 4, pp. 259-267.

Cross J (1995), "IT outsourcing: British Petroleum's competitive approach", *Harvard Business Review*, May-June, pp. 94-102.

Cross J, M Earl, JL Sampler (1997), "Transformation of the IT function at British Petroleum", *MIS Quarterly*, Vol. 21, No. 4, pp. 401-424.

Currie W (1996), "Outsourcing in the private and public sectors: an unpredictable IT strategy", *European Journal of Information Systems*, Vol. 4, pp. 226-236.

Cyert RM and JG March (1963), *A behavioral theory of the firm*, Prentice Hall.

De Loof L (1995), "Information systems outsourcing decision making: a framework, organisational theories and case studies", *Journal of Management Information Systems*, Vol. 10, No. 4, pp. 281-297.

Demsetz H (1993), "The theory of the firm revisited", in Williamson and Winter (1993), pp. 159-178.

Demsetz H (1997) *The economics of the business firm: Seven critical commentaries*, Cambridge University Press.

Donaldson L (1990), "The etheral hand: Organisational economics and management theory", *Academy of Management Review*, Vol. 15, No. 3, pp. 369-381.

Duncan N (1995), "Buying core competencies? A study of the impact of outsourcing in IT infrastructure flexibility", The Inaugural AIS Americas Conference on Information Systems, August 25-27, Pittsburgh, US.

Earl M (1989) *Management strategies for information technology*. Prentice Hall.

Earl M (1996), "The risks of outsourcing IT", *Sloan Management Review*, Spring, pp. 26-32.

Earl M.J., Edwards B. & Feeny D. (1996) Configuring the IS function in complex organisations. In *Information Management The Organisational Dimension*, Earl, M.J. (ed.), pp. 201-230. Oxford University Press.

Eisenhardt KM (1989), "Building theories from case study research", *The Academy of Management Review*, Vol. 14, No. 4, pp. 532-550.

Elden M and RF Chisholm (1993), "Emerging varieties of action research: Introduction to the special issue", *Human Relations*, Vol. 46, No. 2, pp. 121-142.

Fama EF and MC Jensen (1983), "Separation of ownership and control", *Journal of Law and Economics*, Vol. 26, pp. 301-325.

Feeny D, L Willcocks, T Rands and G Fitzgerald (1993), "Strategies for IT management - when outsourcing equals rightsourcing" in Rock S (ed.) *Diectors guide to outsourcing*, Institute of Directors/IBM, London.

Feeny D, M Earl and B Edwards (1994), "Organisational arangements for IS roles of users and IS specialists", *Research and Discussion Papers*, No. 6, Oxford Institute of Information Management.

Fitzgerald G (1994), "The outsourcing of information technology: revenge of the business manager or legitimate strategic option?", Inaugural Lecture, Birkbeck College, University of London, 19 October.

Fitzgerald G and L Willcocks (1994a), "Contracts and partnerships in the outsourcing of IT", *Proceedings of the 15th International Conference on Information Systems*, 14-17 December, Vancouver, Canada, pp. 91-98.

Fitzgerald G and L Willcocks (1994b), "Relationships in outsourcing: contracts and partnerships" in the proceedings, 2nd European Conference on Information Systems, 30-31 May, Nijenrode University, Netherlands, pp. 51-63.

Fitzgerald G, RA Hirschheim, E Mumford and AT Wood-Harper (1985), "Information systems research methodology: An introduction to the debate", in Mumford et al (1985), pp. 3- 9.

Friedberg A and W Yarberry (1994), "Audit rights in an outsource environment", in M Khosrowpour (Ed.), *Managing information technology investments with outsourcing*, Idea Book Publishing, pp. 205-216.

Galbraith JR (1977), *Organisation design*, Addison Wesley.

Galliers RD (1991), "Choosing appropriate information systems research approaches: a revised taxonomy", in HE Nissen, HK Klein and R. Hirschheim (Eds.) *Information systems research: contemporary approaches and emergent traditions*, North Holland, pp. 327-345.

Galliers RD (1995), "A manifesto for information management research", *British Journal of Management*, Vol. 6, Special Issue, pp. s45-s52.

Galliers RD (Ed.) (1992), *Information systems research: issues, methods and practical guidelines*, Blackwell Scientific, Oxford.

Goldberg VP (1976), "Toward an expanded economic theory of contract", *Journal of Economic Issues*, Vol. 10, No. 1, pp. 45-61.

Goldberg VP (Ed.)(1989), *Readings in the economics of contract law*, Cambridge University Press.

Griese J (1993), "Outsourcing of information systems in Switzerland: a status report", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Grossman SJ and OD Hart (1986), "The costs and benefits of ownership: A theory of vertical and lateral integration", *Journal of Political Economy*, Vol. 94, No. 4, pp. 691-719.

Grover V and J Teng (1993), "The decision to outsource information systems functions", *Journal of Systems Management*, November, pp. 34-38.

Grover V, M Cheon and J Teng (1996), "The effect of service quality and partnership on the outsourcing of informaiton systems functions", *Journal of Management Information Systems*, Vol. 12, No. 4, pp. 89-116.

Grover V, M J Cheon and J Teng (1994), "A descriptive study on the outsourcing of information sytems functions", *Information and Management*, Vol 27, pp. 33-44.

Gupta U and A Gupta (1992), "Outsourcing the IS function: is it necessary for your organization?", *Information Systems Management*, Summer, pp. 44-50.

Hart O (1995), *Firms, contracts and financial structure*, Clarendon Press - Oxford.

Hart O and J Moore (1990), "Property rights and the nature of the firm", *Journal of Political Economy*, Vol. 98, No. 6, pp. 1119-1157.

Hart OD (1993), "Incomplete contracts and the theory of the firm", in Williamson and Winter (1993), pp. 138-158.

Heckman R and W King (1994), "Behavioral indicators of customer satisfaction with vendor-provided informaiton services", *Proceedings of the 15th International Conference on Information Systems,* 14-17 December, Vancouver, Canada, pp. 429-444.

Heinzl A (1993), "Outsourcing the information systems function within the company - an empirical survey", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Henderson J (1990), "Plugging into strategic partnerships: the critical IS connection", *Sloan Management Review*, Vol. 31, pp. 7-18.

Hindle T, P Checkland, M Mumford and D Worthington (1995), "Developing a methodology for multidisciplinary action research: A case study", *Journal of Operational Research Society*, Vol. 46, pp. 453-464.

Hodgkinson SL (1996), "The role of the corporate IT function in the federal IT organisation", in M Earl (Ed.), *Information management: The organisational dimension*, Chapter 12. Oxford University Press.

Hodgson GM (1998), "Competence and contract in the theory of the firm", *Journal of Economic Behavior and Organization*, Vol. 35, pp. 179-201.

Holmstrom B (1989), "Agency costs and innovation", *Journal of Economic Behavior and Organization*, Vol. 12, pp. 305-327.

Hopper M (1990), "Rattling SABRE - new ways to compete on information", *Harvard Business Review,* May-June, pp. 118-125.

Huber R (1993), "How continental bank outsourced its crown jewels", Harvard Business Review, January-February, pp. 121-129.

Hufnagel EM and JG Birnberg (1989), "Perceived chargeback system fairness in decentralised organisations: An examination of the issues", *MIS Quarterly*, Vol. 13, No. 4, pp. 415-430.

Hurst I and B Henessian (1995), "Navigating IT channgels: integrate or outsource?", *The McKinsey Quarterly*, No. 3, pp. 103-110.

IDC and WBS (1994), *IT outsourcing: the changing outlook*, International Data Corporation and Information Systems Research Unit, Warwick business School.

Iivari J and I Ervasti (1993), "The impact of alternative IS acquisition strategies upon IS success", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Jensen MC (1983), "Organization theory and methodology", *Accounting Review*, Vol. 50, pp. 319-339.

Jensen MC and WH Meckling (1976), "Theory of the firm: Managerial behavior, agency costs and ownership structure", *Journal of Financial Economics*, Vol. 3, pp. 305-360.

Jones C (1994), "Evaluating software outsourcing options", *Information Systems Management*, Fall, 28-33.

Joskow PL (1993), "Asset specificity and the structure of vertical relationships: empirical evidence", in Williamson and Winter (1993), pp. 117-137.

Jurison J (1995), "The role of risk and return in information technology outsourcing decisions", *Journal of Information Technology*, Vol. 10, No. 4, pp. 239-247.

Kambil A and J Turner (1993), "Information systems outsourcing as organisational alignment and transformation", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Kaplan B and D Duchon (1988), "Combining qualitative and quantitative methods in information systems research: A case study", *MIS Quarterly*, December, pp. 571-586.

Kemmis S (1988) "Action research in retrospect and prospect", in Kemmis and McTaggart (1988), pp. 27-39.

Kemmis S and R McTaggart (1988), *The action research reader*, Third Edition, Deakin University Press.

Kern T and L Willcocks (1996), "The enabling and determining environment: neglected issues in an IS outsourcing strategy", in J D Coelho, T Jelassi, W Konig, H Krcmar, R O'Callaghan and M Saakssjarvi (Eds.), *Proceedings of the 4th European Conference on Information Systems*, Lisbon, Portugal, July 2-4, pp. 1039-1048.

Ketler K and J Walstrom (1993), "The outsourcing decision", Iinternational *Journal of Information Management*, No. 13, pp. 449-459.

Khalil F and Lawarree J (1995), "Input versus output monitoring: who is the residual claimant?", *Journa of Economic Theory*, Vol. 66, pp. 139-157.

Khosrowpour M, G Subramanian and J Gunterman (1994), "Outsourcing: organisational benefits and potential problems" in M Khosrowpour (Ed.), *Managing information technology investments with outsourcing*, in Idea Book Publishing, pp. 244-268.

Klein B (1980), "Transacion cost determinants of 'unfair' contractual arrangements", *American Economic Review*, Vol. 70, No. 2, pp. 356-362

Klein B (1993), "Vertical integration as organisational ownership: the Fisher Body - General Motors relationship revisited", in Williamson and Winter (1993), pp. 213-226.

Klein B, R Crawford and A Alchian (1978), "Vertical integration, appropriable rents and the competitive contracting process", *Journal of Law and Economics*, Vol. 21, pp. 297-326.

Klein HK and K Lyytinen (1985), "The poverty of scientism in information systems", in Mumford et al (1985), pp. 131-162.

Klepper R (1993), "Efficient outsourcing relationships", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Klepper R (1995), "The management of partnerig development in IS outsourcing", *Journal of Information Technology*, Vol. 10, No. 4, pp. 249-258.

Kreps DM (1990), "Corporate culture and economic theory", in JE Alt and KA Shepsle (Eds.) (1990), *Perspectives on positive political economy*, Cabridge University Press, pp. 90-143.

Kriebel C and J Moore (1981), "Economics and management information systems", in *Proceedings of the First Internatioal Conference on Information Systems*, pp. 19-31.

Kumar K and L Willcocks (1996), "Offshore outsourcing: a country too far?", in J D Coelho, T Jelassi, W Konig, H Krcmar, R O'Callaghan and M Saakssjarvi (Eds.), *Proceedings of the 4th European Conference on Information Systems*, Lisbon, Portugal, July 2-4, pp. 1309-1325.

La Belle A and H E Nyce (1987), "Wither the IT organisation?", *Sloan Management Review*, Summer, pp. 75-85.

Lacity M (1992), "An interpretive investigation of the information systems outsourcing phenomenon", PhD Dissertation, University of Houston, USA.

Lacity M and L Willcocks (1995), "Interpreting information technology outsourcing decisions from a transaction cost perspective: findings and critique", *Accounting, Management and Information Technologies*, Vol. 5, No.3/4, pp. 203-244.

Lacity M and R Hirschheim (1993a), "The information systems bandwagon: look before you leap", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Lacity M and R Hirschheim (1993b), *Information Systems Outsourcing: Myths, metaphors and realities*, Wiley Series in Information Systems, Chichester.

Lacity M and R Hirschheim (1994), "IS outsourcing evaluations: lessons from the field", in *Proceedings, IFIP TC8 Information Systems International Working Conference*, Queensland, Australia, May, pp. 363-373.

Lacity M and R Hirschheim (1995), *Beyond the information systems outsourcing bandwagon: the insourcing response*, Chichester: Wiley.

Lacity M C, L Willcocks and D Feeny (1995), "IT outsourcing: maximize flexibility and control", *Harvard Business Review*, May-June, pp. 84-93.

Lacity M, R Hirschheim and L Willcocks (1994), 'Realizing outsourcing expectations: incredible expectations, credible outcomes", *Information Systems Management*, Fall, pp. 7-18.

Lacity M, R Hirschheim and L Willcocks (1995), "Are information systems outsourcing expectations realistic? A review of US and UK evidence", Research and Discussion Paper, 95/1, Oxford Institute of Information Management.

Lacity MC, LP Willcocks and DF Feeny (1996), "The value of selective IT sourcing", *Sloan Management Review*, Spring, pp. 13-25.

Laffont JJ and J Tirole (1993), *A theory of incentives in procurement and regulation*, MIT Press.

Lafontaine F (1992), Ágency theory and franchising: some empirical results", *RAND Journal of Economics*, Vol. 23, No. 2, pp. 263-283.

Lau F (1997), "A review of the use of action research in information systems studies", in AS Lee, J Liebenau and JI DeGross (Eds.), *Information Systems and Qualitative research*, Chapman and Hall, pp. 31-68.

Lawson T (1997), *Economics and reality*, Routledge.

Leibenstein H (1975), "Aspects of the X-efficiency theory of the firm", *Bell Journal of Economics*, Vol. 6, pp. 580-607.

Levitt T (1985), "Paradoxical futures versus a new beginning", in Buzzell RD (Ed.), *Marketing in an electronic age*, Harvard Business School Press: Boston, pp. 364-375.

Lewellyn KN (1931), "What price contract? An essay in perspective", *Yale Law Journal*, Vol. 40, pp. 704-751.

Lieberman MB (1991), "Determinants of vertical integration: An empirical test", The Journal of Industrial Economics, Vol. 39, No. 5, pp. 451-466.

Litwack JM (1993), "Coordination, incentives and the ratchet effect", *RAND Journal of Economics*, Vol. 24, No. 2, pp. 271-.

Loh L (1994), "An organizational-economic blueprint for information technology outsourcing: concepts and evidence", in *Proceedings, 15th International Conference on Information Systems*, 14-17 December, Vancouver, Canada, pp. 73-89.

Loh L and N Venkatraman (1992a), "Diffusion of Information Technology outsourcing: influence sources and the Kodak effect", *Information Systems Research*, Vol. 3, Iss. 4, pp. 334-358.

Loh L and N Venkatraman (1992b), "Determinants of Information Technology outsourcing", *Journal of Management Information Systems*, Vol. 9, No. 1, pp. 7-24.

Loh L and N Venkatraman (1995), "An empirical study of information technology outsourcing: benefits, risks and performance implications", in *Proceedings of the 16th International Conference on Information Systems*, Amsterdam, December 1995, pp. 277-288.

Lowry ST (1976), "Bargain and conract theory in law and economics", *Journal of Economic Issues*, Vol. 10, No. 1, pp. 1-22.

Lundgren K (1990), "Vertical integration, transaction costs and 'learning by using'", in Aoki et al (1990), pp. 112-132.

Lyons BR (1994), "Contracts and specific investment: An empirical test of transaction cost theory", *Journal of Economics and Management Strategy*, Vol.3, No. 2, pp. 257-278.

Lyytinen K (1987), "Different perspectives on information systems: problems and solutions", *ACM Computing Surveys*, Vol. 19, No. 1, pp. 5-46.

Macaulay S (1989), "Non-contractual relations in business: a preliminary study", in Golberg (1989), pp. 4- 15.

Macauley S (1989), "Non-contractual relations in business: a preliminary study", in VP Goldberg, *Readings in the economics of contract law*, Cambridge University Press, pp. 4-15.

Macey A (1993), "Service level agreements and contracts", in S Rock (Ed.), *Director's guide to outsourcing*, Institute of Directors, pp. 48-52.

Macintyre M (1996), "The EDS/Inland Revenue deal two years on", *Business and Technology*, January, pp. 25.

Macneil IR (1978), "Contracts: adjustment of long term economic relationship under classical, neoclassical and relational contract law', *Northwestern University Law Review*, Vol. 72, pp. 854-906.

Macneil IR (1980), *The new social contract*, Yale University Press: New Heaven, CT.

Macneil IR (1985), "Relational contract: What we do and do not know", *Wisconsin Law Review*, pp. 483-525.

Mahoney JT (1992), "The choice of organizational form - vertical financial ownership versus other methods of vertical integration", *Strategic Management Journal*, Vol.13, No.8, pp.559-584.

Main TJ and JE Short (1989), "Managing the merger: Building partnership through IT planning at the new Baxter", *MIS Quarterly*, Vol. 13, No. 4, pp. 469-486.

Markus ML and RI Benjamin (1996), "Change agentry - the next IS frontier", *MIS Quarterly*, Vol. 20, No. 4, pp. 385-408.

Masten SE (1993), "A legal basis for the firm", in Williamson and Winter (1993), pp. 196-212.

Masten SE (Ed.) (1996), *Case studies in contracting and organization*, Oxford University Press.

McFarlan F W and R L Nolan (1995), "How to manage an IT outsourcing alliance", *Sloan Management Review*, Winter, pp. 9-23.

McLellan K, B Marcolin and P Beamish (1995), "Financial and strategic motivations behind IS outsourcing", *Journal of Information Technology,* Vol. 10, No. 4, pp. 299-321.

Mendelson H (1987), "Economies of scale in computing: Grosch's law revisited", *Communications of the ACM,* Vol. 30, No. 12, pp. 1066-1072.

Metcalfe M (1996), *Business research through argument,* Kluwer Academic.

Milgrom P and J Roberts (1990), "Bargaining costs, influence costs and the organisation of economic activity", in JE Alt and KA Shepsle (Eds.) (1990), *Perspectives on positive political economy,* Cabridge University Press, pp. 57-89.

Milgrom P and J Roberts (1990), "The economics of modern manufacturing: techology, strategy and organization", *The American Economic Review,* Vol. 80, pp. 511-528.

Milgrom P and J Roberts (1992), *Economics, organisation and management,* Prentice Hall International.

Milgrom P and J Roberts (1995), "Complimentaries and fit: strategy, structure and organizational change in manufacturing", *Journal of Accounting and Economics,* Vol. 19, pp. 179-208.

Mingers J and F Stowell (Eds.) (1997), *Information systems: an emerging discipline?,* McGraw Hill.

Minoli D (1995), *Analyzing outsourcing,* McGraw-Hill.

Mumford E, R Hirschheim, G. Fitzgerald and T Wood-Harper (Eds.) (1985), *Research methods in information systems,* Proceedings of the IFIP WG 8.2 Colloquium, North Holland.

Muris TJ (1981), "Opportunistic behavior and the law of contracts" *Minnesota Law Review,* Vol. 65, No. 4, pp. 521-590.

Myers MD (1997), "Qualitative research in information systems", *MISQ Discovery Living Scholarship*, http://www.auckland.ac.nz/msis/isworld/ also abstracted in *MISQ Quarterly*, Vol 21, No. 2, pp. 241-242.

Nam K, S Rajagopalan, H Rao and A Chaudhury (1995), "A two stage investigation of the determinants of information technology outsourcing", The Inaugural AIS Americas Conference on Information Systems, August 25-27, Pittsburgh, US.

Nolan R L (1979), "Managing the crises in data processing", *Harvard Business Review*, March-April, pp. 115-126.

Nordhaug O (1998), "Competence specificities in organisations", *International Studies of Management and Organisation*, Vol. 28, No. 1, pp. 8-29.

North DC (1990), *Institutions, institutional change and economic performance*, Cambridge University Press.

Numagami T (1998), "The infeasibility of invariant laws in management studies: a reflective dialogue in defense of case studies", *Organisation Science*, Vol. 9, No. 1, pp. 2-15.

Orlikowski WJ and JJ Baroudi (1991), "Studying information technology in organisations: research approaches and assumptions", *Information Systems Research*, Vol. 2, No. 1, pp. 1- 28.

Ouchi WG (1980), "Markets, bureaucracies and clans", *Administrative Science Quarterly*, Vol. 25, pp. 129-141.

Outhwaite W (1983) *Toward a realist perspective*, in G. Morgan (Ed.) *Beyond method, strategies for social research*, Sage.

Outhwaite W (1987) *New philosophies of social science*, St. Martin's Press.

Panday K (1995), "Clients fine tune FM contracts", *Management Consultancy*, April, pp. 43-44.

Panzar JC and RD Willig (1981), "Economies of scope", *American Economic Review*, Vol. 71, No. 2, pp. 268-272.

Patane J and J Jurison (1994), "IS global outsourcing diminishing prospects for american programmers?", *Journal of Systems Management*, Vol. 45, pp. 6-10.

Peak D (1994), "The risks and effects of outsourcing on the IS function and the firm", PhD thesis, University of North Texas.

Peak D and J Windsor (1993), "The risks and effects of outsourcing on the IS function and the firm", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Pervan GP and Klass DJ (1992), "The use and misuse of statistical methods in information systems research", in Galliers (1992).

Pettigrew AM (1985), "Contextualist research and the study of organisational change processes", in Mumford et al (1985), pp. 53-78.

Pfeffer J (1997), *New directions for organization theory, problems and prospects*, Oxford University Press.

Pfeffer J and GR Salancik (1978), *The external control of organisations: A resource dependence perspective*, Harper and Row.

Pinnington A and P Woolcock (1994), "How far is IS/IT outsourcing enabling new organisational structure and competences?", *in Proceedings, Second SISNET Conference*, Barcelona, Spain, 26-28 September.

Pitt LF, RT Watson and CB Kavan (1995), "Service quality: a measure of information systems effectiveness", *MIS Quarterly*, Vol. 19, No. 2, pp. 173-188.

Poole MS and AH Van de Ven (1989), "Using paradox to build management and organisation theories", *The Academy of Management Review*, Vol. 14, No. 4, pp. 562-578.

Pratt JW and RJ Zeckhauser (1985), *Principals and agents: the structure of business*, Harvard Business School Press, Boston.

Putrus R (1994), "Outsourcing and justification using AHP", in M Khosrowpour (Ed.), *Managing information technology investments with outsourcing*, Idea Book Publishing, pp. 159-204.

Putterman L (1995), "Markets, hierarchies and information: On a paradox in the economics of organisation", *Journal of Economic Behavior and Organization*, Vol. 26, pp. 373-390.

Quinn R and J Rohrbaugh (1983), "A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis", *Management Science*, Vol. 29, No. 3, pp. 363-377.

Rands T (1992), "The key role of applications software make-or-buy decisions", *Journal of Strategic Information Systems*, Vol. 1, No. 4, pp. 215-223.

Rapoport RN (1970), "Three dilemmas in action research", *Human Relations*, Vol. 23, No. 6, pp. 499-513.

Reve T (1990), "The firm as a nexus of internal and external contracts", in M Aoki, B Gustaffosn and O Williamson (Eds.), *The firm as a nexus of treaties*, Newbury Park, Sage.

Richmond W and A Seidmann (1993), "Software development outsourcing contract: structure and business value", *Journal of Management Information Systems*, Vol. 10, No. 1, pp. 57-72.

Richmond W, A Seidmann and A Whinston (1992), "Incomplete contracting issues in information systems development outsourcing", *Decision Support Systems*, Vol. 8, No. 5, pp. 459-477.

Rindfleisch A and JB Heide (1997), "Transaction cost analysis: part, present and future applications", *Journal of Marketing*, Vol. 61, October, pp. 30-54.

Riordan M and OE Williamson (1985), Ásset specificity and economic organization", *International Journal of Industrial Organization*, Vol. 3, pp. 365-378.

Riordan MH (1990), "What is vertical integration?", in Aoki et al (1990), pp. 94-111.

Riordan MH and DEM Sappington (1987), "Information, incentives and organizational mode", *Quarterly Journal of Economics*, Vol. 52, pp. 243-263.

Roberts PW and R Greenwood (1997), "Integrating transaction cost and institutional theories: toward a constrained-efficiency framework for understanding organisational design adoption", *Academy of Management Review*, Vol. 22, No. 2, pp. 346-373.

Rogerson WP (1992), "Contractual solutions to the hold-up problem", *Review of Economic Studies*, Vol. 59, pp. 777-794.

Rosenberg B and C Macaulay (1993) *Facilities management / outsourcing: a strategic business function for the 1990s*, IBC Publishing.

Rothery B and I Robertson (1995) *The truth about outsourcing*, Gower.

Saaksjarvi M (1993), "Outsourcing of information systems: matching organisational forms and IS roles", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Saarinen T and A Vepsalainen (1994), "Procurement strategies for information systems", *Journal of Management Information Systems*, Vol. 11, No. 2, pp. 187-208.

Saarinen T and M Saaksjarvi (1993), "Empirical evaluation of two different IS outsourcing strategies in Finish wood working industries", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

Sandberg A (1985), "Socio-technical design, trade union strategies and action research", in Mumford et al (1985), pp. 79-92.

Sappington DEM (1991), "Incentives in principal-agent relationships", *Journal of Economic Perspectives*, Vol. 5, No. 2, pp. 45-66.

Sayer A (1992), *Method in social science: A realist approach*, Second Edition, Routledge.

Sen A (1996), "Termination clauses in long term contracts", *Journal of Economics and Management Strategy*, Vol. 5, No. 4, pp. 473-496.

Sen A (1996), "Termination clauses in long-term contracts", *Journal of Economics and Management Strategy*, Vol. 5, No. 4, pp. 473-496.

Silverman D (1998), "Qualitative research: meanings or practices", *Information Systems Journal*, Vol.8, pp. 3-20.

Simon H, M Edigi, R Marris and R Viale (Eds.) (1992), *Economics, bounded rationality and the cognitive revolution*, Edward Elgar.

Simon HA (1961), *Administrative behavior*, Macmillan.

Simon HA (1991), "Organizations and markets", *Journal of Economic Perspectives*, Vol. 5, No. 2, pp. 25-44.

Smith HJ, SJ Milberg and SJ Burke (1996), "Information privacy: measuring individual's concerns about organisational practices", *MIS Quarterly*, Vol. 20, No. 2, pp. 167-196.

Sobol M and U Apte (195), "Domestic and global outsourcing practices of America's most effective IT users", *Journal of Management Information Systems*, Vol. 10, No. 4, pp. 269-280.

Stigler G (1968), "The division of labor is limited by the extent of the market", in Stigler G, *The organisation of industry*, University of Chicago Press, pp. 129-141.

Strowger B (1993), "The intelligent customer", in S Rock (Ed.), *Director's guide to outsourcing*, Institute of Directors, pp.60-64.

Szymankiewicz J (1993), *Contracting out or selling out? Survey into the current issues concerning the outsourcing of distribution*, P-E International.

Takac P (1994), "Outsourcing: a key to controlling escallating IT costs?", *International Journal of Technology Management*, Vol. 9, No. 2, pp. 139-155.

Tavakolian H (1989), "Linking the information technology structure with organisational competitive strategy: A survey", *MIS Quarterly*, Vol. 13, No. 3, pp. 309-317.

Taylor JR and CC Tucker (1989), "Reducing data processing costs through centralised procurement", *MIS Quarterly*, Vol. 13, No. 4, pp. 487-499.

Teece DJ (1980), "Economies of scope and the scope of the enterprise", *Journal of Economic Behaviour and Organization*, Vol. 1, pp. 223-247.

Teng J, M Cheon and V Grover (1995), "Decisions to outsource information systems functions: testing a strategy-theoretic discrepancy model", *Decision Sciences*, Vol. 26, No. 1, pp. 75-103.

Thong J, C S Yap and K S Raman (1995), "Engagement of external expertise in information systems implementation", *Journal of Management Information Systems*, Vol. 1, No. 2, pp. 209-231.

Tirole J (1986), "Hierarchies and bureaucracies: On the role of collusion in organisations", *Journal of Law, Economics and Organization*, Vol. 2, No. 2, pp. 181-214.

Tsoukas H (1989), "The validity of idiographic research explanations", *Academy of Management Review*, Vol. 14, No. 4, pp. 551-561.

Ulset S. (1996), "R&D outsourcing and contractual governance: an empirical study of commercial R&D projects", *Journal of Economic Behavior and Organization*, Vol. 30, pp. 63-82.

Umanath NS, MR Ray and TL Campbell (1996), "The effect of unertainty and information asymmetry on the structure of compensation contracts", *Maanagement Science*, Vol. 42, No. 6, pp. 868-874.

Valor J, D Fonstad and R Andreu (1993), "Outsourcing in Spain: An empirical study of top management's perspective", Research paper No 244, March, Research Division, IESE, University of Navarra, Barcelona, Spain.

Van Alstyne M, E Brynjolfddon and S Madnick (1995), "Why not one big database? Principles for data ownership", *Decision Support Systems*, Vol. 15, pp. 267-284.

Venkatraman N and L Loh (1994), "The shifting logic of the IS organization: from technical protfolio to relationship portfolio", *Information Strategy: The Executive's Journal*, Winter, pp. 5-11.

Von Simson EM (1990), "The 'centrally decentralised' IS organisation", *Harvard Business Review*, July-August, pp. 158-162.

Vowler J (1996a), "Braveheart", *Computer Weekly*, 28 November, pp.48-49.

Vowler J (1996b), "Scots take the outside road to IT warehouse", *Computerweekly*, 5 September 1996, pp. 1.

Walsham G (1995), "The emergence of interpretivism in IS research", *Information Systems Research*, Vol. 6, No. 4, pp. 376-394.

Wang E and T Barron (1995a), "The decision to outsource IS processing under internal information asymmetry and conflicting objectives", *Journal of Organizational Computing*, Vol. 5, No. 3, pp. 219-253.

Wang E and T Barron (1995b), "Controlling information system departments in the presence of cost information asymmetry", *Information Systems Research*, Vol. 6, No. 1, pp. 24-50.

Ward J (1995) *Principles of information systems management.* Routledge.

Ward J and J Peppard (1996), "Reconciling the IT/business relationship - a troubled marriage in need of guidance", *Journal of Strategic Information Systems*, Vol.5, No.1, pp.37-65.

Ward S and C Chapman (1994), "Choosing contractor payment terms", *International Journal of Project Management*, Vol. 12, No. 4, pp. 216-221.

Wassenaar A and T Thiadens (1993), "Outsourcing or insourcing? That will be the question - A transactional view on Dutch experiences of outsourcing and insourcing", Outsourcing Information Systems Services Conference, University of Twente, Netherlands, May 20-22.

West L (1994), "Researching the costs of information systems", *Journal of Management Information Systems*, Vol. 11, No. 2, pp. 75-107.

Whang S (1992), "Contracting for software development", *Management Science*, Vol. 38, No. 3, pp. 307-324.

Whang S (1995), "Market provision of custom software: learning effects and low balling", *Management Science*, Vol. 41, No. 8, August, pp. 1343-1352.

Whang S and T Barron (1995), "The decision to outsource IS processing under internal information asymmetry and conflicting objectives", *Journal of Organizational Computing*, Vol. 5, No. 3, pp. 219-253.

Whitford WC (1985), "Ian Macneil's contribution to contracts scholarship", *Wisconsin Law Review*, pp. 545-560.

Willcocks L (1994), "Collaborating to compete: towards strategic parnterships in IT outsourcing?", *Research and Discussion Papers*, No. 11, Oxford Intitute of Information Management.

Willcocks L and Choi C J (1995), "Co-operative partnership and 'total' IT outsourcing: from contractual obligation to strategic alliance?", *European Management Journal*, Vol. 13, No. 1, pp. 67-78.

Willcocks L and G Fitzgerald (1993a), "IT outsourcing: preliminary findings from recent UK research", *Research and Discussion Paper*, No. 10, Oxford Institute of Information Management.

Willcocks L and G Fitzgerald (1993b), "Market as opportunity? Case studies in outsourcing information technology and services", *Journal of Strategic Information Systems*, Vol. 2, No. 3, September 1993, pp. 223-242.

Willcocks L and G Fitzgerald (1994a), "IT outsourcing and the changing shape of the information systems function: recent research findings", *Research and Discussion Paper,* No. 4, Oxford Intitute for Information Management.

Willcocks L and G Fitzgerald (1994b), "To outsource IT or not? Recent research on economics and evaluation practice", in A Brown and D Remenyi (Eds.), *Proceedings of the First European Conference on IT Investment Evaluation,* Henley on Thames, England, 13-14 September, pp. 235-247.

Willcocks L and G Fitzgerald (1994c), *A business guide to outsourcing IT,* Business Intelligence.

Willcocks L and G Fitzgerald (1994d), "Toward the residual IS organisation? Research on IT out experiences in the United Kingdom", in R Baskerville, S Smithson, O Ngwenyama and J I DeGross (Eds.), *Transforming organizations with information technology,* Elsevier Science, IFIP.

Willcocks L and M Lacity (1995), "Information systems outsourcing in theory and practice", *Journal of Information Technology,* Vol. 10, No. 4, pp. 203-207.

Willcocks L, G Fitzgerald and D Feeny (1994), "Information technology outsourcing: from incrementalism to strategic intent", *in Proceedings, Second SISNET Conference,* Barcelona, Spain, 26-28 September.

Willcocks L, M Lacity and G Fitzgerald (1995), "IT outsourcing in Europe and the USA: assessment issues", in G Doukidis, B Galliers, T Jelassi, H Krcmar and F Land (Eds.), *Proceedings of the 3rd European Conference on Information Systems,* Athens, Greece, 1-3 June. pp. 247-260.

Willcocks LP , G Fitzgerald and M Lacity (1996), "To outsource IT or not? recent research on economics and evaluation practice", *European Journal of Information Systems,* Vol. 5, pp. 143-160.

Willcocks LP and MC Lacity (1998), *Strategic sourcing of information systems: Perspectives and practices,* Wiley Series in Information Systems, Chichester.

Williamson OE (1975), *Markets and hierarchies: Analysis and antitrust implications,* Free Press.

Williamson OE (1985), *The economic institutions of capitalism*, Free Press, NY.

Williamson OE (1986), "Transaction cost economics: the governance of contractual relations", in OE Williamson (Ed.), *Economic Organisation*, Whitesheaf: Brighton, pp. 101-130.

Williamson OE (1990), *Organization theory: from Chester Barnard to the present and beyond*, Oxford University Press, NY.

Williamson OE (1991), "Comparative economic organisation: the analysis of discrete structural alternatives", *Administrative Science Quarterly*, Vol. 36, pp. 269-296.

Williamson OE (1993), "Calculativeness, trust and economic organisation", *Journal of Law and Economics*, Vol. 36, pp. 453-486.

Williamson OE (1996), *The mechanisms of governance*, Oxford University Press.

Williamson OE and SG Winter (Eds.) (1993), *The theory of the firm: Origins, evolutions and development*, Oxford University Press.

Winder M (1994), "Transitional outsourcing", *Information Systems Management*, Fall, pp. 65-68.

Winter SG (1993), "On Coase, competence and the corporation", in Williamson and Winter (1993), pp. 179-195.

Wood-Harper AT (1985), "Research methods in information systems: using action research", in Mumford et al (1985), pp. 169-192.

Yarbrough BV and RM Yarbrough (1988), "The transactional structure of the firm: A comparative survey", *Journal of Economic Behavior and Organization*, Vol. 10, pp.1-28.

Yin RK (1989), *Case study research: design and methods*, Revised Edition, Sage.

Zaheer A and N Venkatraman (1995), "Relational govenrance as an interorganizational strategy: an empirical test of the role of trust in economic exchange", *Strategic Management Journal*, Vol. 16, pp. 373-392.

Zmud R (1984), "Design alternatives for organizing information systems activities", *MIS Quarterly*, June, pp. 79-93.

Zmud R (1996), "Editor's comments", *MIS Quarterly*, Vol. 20, No. 2, pp. xxi-xxiv.

Zmud R, AC Boynton and GC Jacobs (1986), "The information economy: A new perspective for effective information systems management", *Data Base*, Vol. 18, No. 1, pp. 17-23.

APPENDIX 1 Interview Guide

The following was used as a guide to the interviews carried out at ICI andAnglian Water. It has been derived from the S-CAGE framework as presented in chapter 5.

Plan

- Interview Part A: Discussing the overall organisation of IT and the allocation of services to contracts. *Likely contact:* IT director.
- Interview Part B: Discussing the governance details of selected contracts. While specific services are the focal points, the whole corresponding contract will be examined. *Likely contact:* Contract manager or equivalent; someone who knows the contract well.

Interview guide

{The corresponding elements of S-CAGE are indicated in curly brackets}

Part A

{Service clustering}

- Obtain a comprehensive list and/or statistics of the kinds of contracts (services, duration, value, compensation, management involvement) should be recorded.
- Discuss why have they grouped/split services in the way they have and not in some other way. Discuss any plans for changes.

Part B

{General}

- Start date and duration of the contract
- Brief history of the outsourcing decision and process
- Discuss success/satisfaction with the contract.

{1. Assurance of service delivery}

{1.1. Service specification}

- List the services and activities identified in the contract.
- Discuss why they were put under the same contract. What do they have in common? What are the most profound differences?
- Discuss the degree of specificity of assets or investments included in the contract (physical (h/w, s/w, data), human, site and dedicated). How important are specific assets in comparison to the total value of the contract?
- Discuss the extent of uncertainty involved with each of the services of the contract. Uncertainty can be defined in three main ways: (I) Changing user needs and requirements, (ii) changes in the technology and (iii) Unpredictable factors affecting performance.
- Is the service or its requirements expected to change over time? How would the contract cope with any changes?
- Discuss the size of operations required and potential economies of scale.
- Discuss possible resources shared among the services within the contract or with other contracts (H/W, S/W, physical space, skills).
- Is the performance of the focal service related to the performance of other services?
- Are any other inter-dependencies and complementarities among services within the same contract or between contracts perceived significant?
- Discuss variations in demand for the service and resource utilisation over time.

- To what extent does the contract specify *what* (the user's view of the service) must be delivered and *how* (e.g. resources, people, skills, methods etc.) it must be done? Why?
- How much discretion and control does the contract give the client to tell the provider *how* to provide the service? Why?
- How critical is this service to the company? What are its most critical performance aspects?

{1.2. Service level requirements specification}

- Discuss how rigorously performance metrics and targets are specified. Both for inputs (costs and perhaps human resources) and for outputs (service levels).
- To what extend are they expressed in business or technical terms.
- Discuss the completeness of SLRs, i.e. the extent to which they cover all aspects of intended performance.
- Discuss the precision of SLRs, i.e. the extent to which they are accurate and unambiguous measures of actual performance.
- Discuss problems that have been observed with respect to the efficacy of SLRs and the actions taken to alleviate them.
- What, if any, different categories of SLRs are there (e.g. critical, non-critical, random)?

{1.3. Service level measurement and verification}

- What monitoring, measurement and reporting mechanisms are specified in the contract, i.e. how are SLRs measured and reported?
- Who is responsible for measurement and reporting?
- Any verification processes regularly carried out or exceptionally provided for by the contract?
- Is there any additional information over and above the SLRs being measured, e.g. persistent problems or exceptions.

- Is the administrative cost of measurement and verification considered to be a problem? Are the measurement/monitoring processes considered to be effective?
- What is the frequency of performance measurement and reporting?
- Have they observed problems with the provider(s) taking advantage of either gaps in monitoring or over-monitoring?
- Have there been disagreements between client and provider as to the true performance delivered?

{2. Incentives for service level attainment}

{2.1. Implicit incentive patterns}

- How difficult is it to obtain information about the costs, capabilities and performance of the supplier?
- Does the contract include the right to audit and/or to benchmark the supplier's activity? Do they make use of this right?
- Have they observed opportunistic behaviour from the provider's part? Does this opportunistic behaviour consist of major failures to meet agreed targets or does it consist of taking advantage of minor gaps in the contract or in the monitoring process?
- How do they react to such behaviour?
- What is the provider's liability in case of failure?
- Who has ownership and who has control rights over hardware and software?
- Who keeps the hardware and software after the end of the contract?
- Are there equivalent provisions for staff to return to the company?
- Are there any exclusivity or confidentiality clauses as to the use of specialised systems or equipment?
- How important is the supplier's reputation, within the particular contract and within the whole market?

{2.2. Contractual reward/penalty devices}

- What proportion of the supplier's total compensation is fixed, what proportion is cost-based and what proportion is based on performance incentives?
- Discuss any explicit incentives stipulated in the contract.
- Does the incentive system encourage by rewarding or does it deter by penalising, or both? Why?
- How does the incentive mechanism work? Is the reward/penalty fixed in advance or is it variable depending on the extent of good/bad performance?
- Does it cover all services under the contract? Does it place more weight on some services than on others?
- Are incentives based on the outcome of the provider's work (i.e. service levels) or on their effort (i.e. cost)?
- Are the incentive targets specified in terms of values, intervals or thresholds?
- Are incentives based on average performance over time or not?
- Are uncontrollable factors affecting the performance of the supplier taken into account?
- Does the incentive system forgive some deviations from the targets or not and how?
- Are incentive targets subject to renegotiation? Have they been revised at all? Are they considered sufficient?
- Are they satisfied with the operation of the incentive system? Have there been any problems?

{2.3. Other contractual and organisational safeguards}

- What, in their opinion, is the strongest incentive for the supplier?
- How important is trust and organisational culture?
- What types of disagreements/problems/disputes are most common? How severe are they perceived to be? How are they dealt with?
- Is there a termination clause and how can it be invoked?

{3. Organisation and management of service provision}

{3.1. Roles and responsibilities}

- What managerial roles does the company and the provider put in place to oversee service delivery and contract execution?
- Discuss the extent and the tasks of the managerial involvement of the company in overseeing the contract.
- How are renegotiations dealt with?

{3.2. Coordination and communication mechanisms}

- Discuss the communication channels that exist between client and provider regarding contract execution and service delivery. Are they formally stipulated in the contract or not?
- How are contingencies dealt with?
- Is there an escalation process? How is it invoked?

{3.3. Locus and scope of decision making}

- Discuss which managers or committees have decision authority over the contract and the relations with the vendor.
- Discuss how disputes are being handled.
- Discuss how renegotiations can be initiated and who participates.

APPENDIX 2 Research Participants

This appendix lists the managers who were involved in interviews and other meetings during the case studies.

BP

The primary contact throughout the case study was David Beaney, then Global Head of IT, BP Oil. Extensive and repeated interaction took place with his associates in Oil's global IS group, Brian Ralphs and Stephen Cotten, during most of the case study (both moved from that office before the end of the case study). Paul Stone, who was initially coordinating IT operations at the Corporate Centre, was involved in the case study from the first presentation. He became much more actively involved in the case study after the departure of Stephen Cotten, when his responsibilities were also augmented. The other attendants at the first presentation were John Cross, Jim Hood (both from BPX) and Graham Williams (BPCC). Most of them were individually interviewed in subsequent meetings. More brief discussions took place with John Cross and Stephen Giles (Oil Europe). During another visit, Tom Gallen (BPCC) offered more extensive information about the IT contracting practices across the Group. Brian Hitch (VP Operations, BP America) offered his view during a long meeting. George Fish (BPX) extensively explained the BPX approach. John Maher (PRM at Oil's HQ at Hemel) clarified the SLAs and grief point systems. Barry Poulten (Air BP at Hemel) explained his role as part time PRM and Air's contract. Robert Booth (SEMA Site Manager at Hemel) and David Tait (SEMA Account Manager, based at Birmingham) offered the supplier's view.

ICI

One interview at the London Headquarters and several telephone interviews were conducted with Dr. Richard Sykes, Group IS Manager.

Anglian Water

Mrs. Diane Rafter, Service Manager, offered a face-to-face and a telephone interview, and a review of the case write up.